

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:22:32 ; Search time 31.5 Seconds  
(without alignments)  
3375.248 Million cell updates/sec

Title: US-09-509-994-1

Perfect score: 2916

Sequence: 1 MGVLVGLALAGLGFAP.....PSPTPGSLTPPAVGLVHSG 516

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_21.\*

1: sp\_archea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_prodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriapi.\*  
17: sp\_archeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result NO.	Score	% Match	Query Length	DB ID	Description
1	2654	91.0	468	4 Q9UC32	Q9uc32 homo sapien
2	1840	63.1	577	11 Q35370	Q35370 rattus norv
3	1396.5	47.9	461	11 P97883	P97883 rattus norv
4	580.5	19.9	757	4 Q9HCU0	Q9hcu0 homo sapien
5	575	19.7	765	11 Q912V1	Q912v1 mus musculu
6	575	19.7	765	11 Q91V98	Q91v98 mus musculu
7	433	14.8	1664	5 Q9TVQ2	Q9tvq2 caenorhabdi
8	414	14.2	1574	11 Q88281	Q88281 rattus norv
9	385	13.2	1394	5 Q9V589	Q9v589 drosophila
10	375	12.9	708	13 P87363	P87363 gallus gall
11	374	12.8	2906	11 Q9WUH9	Q9wuh9 rattus norv
12	369	12.7	1246	4 Q75095	Q75095 homo sapien
13	366	12.6	3857	11 Q88840	Q88840 mus musculu
14	365	12.5	2809	4 Q96JP8	Q96jp8 homo sapien
15	360	12.3	528	11 Q9CXD8	Q9cxd8 mus musculu
16	360	12.3	2872	11 Q9WUH8	Q9wuh8 rattus norv

```

17 359 12.3 937 4 Q96FT5 Q96ft5 homo sapien
18 358.5 12.3 1174 11 Q99K58 Q99k58 mus musculu
19 353.5 12.1 746 4 Q96HB9 Q96hb9 homo sapien
20 353.5 12.1 1256 4 Q9NS15 Q9ns15 homo sapien
21 353.5 12.1 1382 4 Q9H7K2 Q9h7k2 homo sapien
22 352.5 12.1 741 4 Q96K89 Q96k89 homo sapien
23 351.5 12.1 1511 4 Q96K89 Q96k89 homo sapien
24 351.5 12.1 1587 4 Q96RW7 Q96rw7 homo sapien
25 343 11.8 5636 4 Q96RW7 Q96rw7 homo sapien
26 341 11.7 576 4 Q9Y3V7 Q9y3v7 homo sapien
27 338 11.6 1062 11 Q60789 Q60789 mus musculu
28 337.5 11.6 1289 5 Q8SSS3 Q8ss3 dictyosteli
29 337 11.6 2673 4 Q96SC3 Q96sc3 homo sapien
30 333.5 11.4 961 11 Q9EOC6 Q9eqc6 mus musculu
31 333 11.4 999 4 Q9NQ36 Q9nq36 homo sapien
32 331.5 11.4 956 11 Q8R542 Q8r542 mus musculu
33 331 11.4 589 5 Q9TZS1 Q9tzzs1 caenorhabdi
34 330.5 11.3 956 11 Q99K64 Q99k64 mus musculu
35 330 11.3 2189 5 Q9BK05 Q9bk05 elmeria ten
36 328.5 11.3 1713 11 Q88349 Q88349 mus musculu
37 328.5 11.3 1764 11 Q35806 Q35806 rattus norv
38 327.5 11.2 1095 11 Q60784 Q60784 mus musculu
39 323 11.1 1253 11 Q61810 Q61810 mus musculu
40 320.5 11.0 937 5 Q9BLJ1 Q9blj1 ciona intes
41 320 11.0 1394 4 Q8TD95 Q8td95 homo sapien
42 319.5 11.0 1833 11 Q08999 Q08999 mus musculu
43 315.5 10.8 443 4 Q96TF5 Q96tf5 homo sapien
44 315 10.8 683 4 Q8TBH8 Q8tbh8 homo sapien
45 315 10.8 997 11 Q9JJ50 Q9jj50 mus musculu

```

#### ALIGNMENTS

#### RESULT 1

```

Q9UC32 PRELIMINARY; PRT; 468 AA.
ID Q9UC32
AC Q9UC32;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE THROMBOMODULIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=93293792; PubMed=8390446;
RA Yamamoto S., Mizoguchi T., Tamaki T., Ohkuchi M., Kimura S., Aoki N.;
RT "Urinary thrombomodulin, its isolation and characterization.";
RL J. Biochem. 113:433-440(1993).
DR HSSP; P07204; 1ZAO.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF-Ca.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR001491; Thrombomoduln.
DR Pfam; PF00008; EGF_4.
DR PRINTS; PR00059; lectin_c; 1.
DR SMART; SM00034; CLECK; 1.
DR SMART; SM00181; EGF; 6.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR PROSITE; PS00041; C-TYPE-LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01187; EGF_CA; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 468 AA; 49444 MW; 4BFE8E98EFB6A40 CRC64;

```

Query Match 91.08; Score 2654; DB 4; Length 468;  
Best local similarity 99.6%; Pred. No. 7.2e-205;  
Matches 466; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 19 APAEPQGGSCQVEHDCFCALYPGPATFLNASQICDGLRHMTVRSVAADVLSLLNGD 78
DB 1 APAEPQGGSCQVEHDCFCALYPGPATFLNASQICDGLRHMTVRSVAADVLSLLNGD 60
QY 79 GVGRRRLWIGLQPLPGCGDPRKLRGLRQFWWTGDNNTYSRWARLDLNGAPLCPCLV 138
DB 61 GVGRRRLWIGLQPLPGCGDPRKLRGLRQFWWTGDNNTYSRWARLDLNGAPLCPCLV 120
QY 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEFEHPATCRPLAVPGAAAVSITYTTPFA 198
DB 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEFEHPATCRPLAVPGAAAVSITYTTPFA 180
QY 199 ARGADQALPVSSAAVAPLGLQLMCTAPPAGVQGHWAREAPGAWDCSVENGCEHACNA 258
DB 181 ARGADQALPVSSAAVAPLGLQLMCTAPPAGVQGHWAREAPGAWDCSVENGCEHACNA 240
QY 259 IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNPDPQGSYSCMCTGYRLAA 318
DB 241 IPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNPDPQGSYSCMCTGYRLAA 300
QY 319 DQHRCEVDVDCILERSPCQRCVNTQGGFECCHYPNYDLVGECEVPDPCFRANCEYQC 378
DB 301 DQHRCEVDVDCILERSPCQRCVNTQGGFECCHYPNYDLVGECEVPDPCFRANCEYQC 360
QY 379 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQACPADCDPNTQASCEGPEGYILDDGFI 438
DB 361 QPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQACPADCDPNTQASCEGPEGYILDDGFI 420
QY 439 CTDIDECENGFGSCGVCHNLPGTFECICGPDPSALVRHIGTDCDSKVD 486
DB 421 CTDIDECENGFGSCGVCHNLPGTFECICGPDPSALVRHIGTDCDSKVD 468

RESULT 2
O35370 PRELIMINARY; PRT; 577 AA.
AC O35370;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE THROMBOMODULIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;
RA Wang J., Yao A., Wang J.Y., Hardin J.W., Hauer-Jensen M.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 1-136 FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Yao A., Wang J., Hardin J.W., Hauer-Jensen M.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF022743; AAB80760.1; -.
DR EMBL; AF022742; AAB80923.1; -.
DR HSSP; P07204; 1FGD.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF-Ca.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR001491; Thrbomodulin.
DR Pfam; PF00008; EGF; 5.
DR Pfam; PF00059; lectin_c; 1.
DR PRINTS; PR00907; THRBOMODULN.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00001; EGF_like; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS00041; C_TYPE_LLECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 3.
```

```
DR PROSITE; PS01187; EGF_CA; 2.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 577 AA; 61844 MW; 0BE764C8BF18555F CRC64;

Query Match 63.1%; Score 1840; DB 11; Length 577;
Best Local Similarity 63.8%; Pred. No. 1.8e-139;
Matches 332; Conservative 49; Mismatches 133; Indels 6; Gaps 5;

QY 1 MLGVLVIGLALAGLFPAPAEQPGGSCQVEHDCFCALYPGPATFLNASQICDGLRHLM 60
DB 1 MLGVLVIGLALAGLFPAPAEQPGGSCQVEHDCFCALYPGPATFLNASQICDGLRHLM 60
QY 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPRKLRGLRQFWWTGDNNTSYS 120
DB 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPRKLRGLRQFWWTGDNNTSYS 118
QY 121 RWARLDLNGAPLCPCLVCAVSAAEATVPSEPIWEEQCEVKADGFLCEFEHPATCRPLAV 180
DB 119 RWARLDLNGAPLCPCLVCAVSAAEATVPSEPIWEEQCEVKADGFLCEFEHPATCRPLAV 178
QY 181 EP-GAAAAAVSITYTTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREA 239
DB 179 NTRDPEGAHSSITYNTPLVSGADFTLPIGSSATVAPFGLVLCRALPGTSGEHWTRV 238
QY 240 PGAWDCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 299
DB 239 TGAWDCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 298
QY 300 PDQPGSYSCMCTGYRLAADQHRCEVDVDCILERSPCQRCVNTQGGFECCHYPNYDLVD 359
DB 299 SDVPGSYSCMCTGYRLAADQHRCEVDVDCILERSPCQRCVNTQGGFECCHYPNYDLVD 358
QY 360 GECVEPVDPCFRANCEYQCPNLQTSYLCVCAEGFAPIPHEPHRCQMFNCQACPADCDP 419
DB 359 GECVEPVDPCFRANCEYQCPNLQTSYLCVCAEGFAPIPHEPHRCQMFNCQACPADCDP 418
QY 420 NTOASCECPGYILDDGFICTDIDECENGFGSCGVCHNLPGTFECICGPDPSALVRHIGT 479
DB 419 NSPFCOCPEGFILDESGICTDIDECENGFGSCGVCHNLPGTFECICGPDPSALVRHIGT 478
QY 480 CDSGKV-DGGDSGSGSEPPPS-PTPGSTLTPPAVLGVHSG 516
DB 479 CDPVLEDESDGSGSEHPSSNPTVSSVTPPSARPMHSG 518

RESULT 3
P97883 PRELIMINARY; PRT; 461 AA.
ID P97883;
AC P97883;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Thrombomodulin (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRIN CAPILLARY;
RA Wang L., Tran N.D., Schreiber S.S., Zlokovic B.V.;
RL Nucleotide sequence of rat thrombomodulin.;
DR EMBL; U90121; AAB49723.1; -.
DR HSSP; P07204; 1FGD.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF-Ca.
DR InterPro; IPR001491; Thrbomodulin.
DR Pfam; PF00008; EGF; 5.
DR PRINTS; PR00907; THRBOMODULN.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00001; EGF_like; 3.
```

```
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS01186; EGF_2; 3.
DR PROSITE; PS01187; EGF_CA; 2.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
FT NON_TER 1 1
SQ SEQUENCE 461 AA; 49499 MW; 45971BCB84688B67 CRC64;

Query Match 47.9%; Score 1396.5; DB 11; Length 461;
Best Local Similarity 62.4%; Pred. No. 5.2e-104;
Matches 244; Conservative 37; Mismatches 105; Indels 5; Gaps 4;

Qy 131 PLGGLPLVAVSAARATVPSEIWEQCEVKADGFLCEFFHPATC-RPLAVEP-GAAAAA 188
Db 12 PLGGLPLVTVSTATAPGEPFAWEKFCENETKGLCEFFYFAAFCSRLRYNTRDPEGAH 71
Qy 189 VSIYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVGOVHWAREAPGAWDCSVE 248
Db 72 ISSYNTPLVGVGADFQTLPGSSATVAPFGLVLCRALPDTSGHWTREVTGAWNCSVE 131
Qy 249 NGGCEHACNAPLGPAPCOPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPQGSYSY 308
Db 132 NGGCEYMCNRSANGPRCVPSCGGDLQADGRSCAKPVGLCNELCOHFCVNNSDVPGSYSY 191
Qy 309 MCEGTGYRLAADQHRCEVDVDDCILPSPQPCQRCVNTQGGFCHCYPNLDLVDCGEVPEVDP 368
Db 192 MCEGTGYRLAADQHRCEVDVDDCKQGNPCQPLCSNTEGFCRCYDGYELVDGECVEQLDP 251
Qy 369 CFRANCEYQCOPLQNTSYLVCAEGEAFIPHEPHRCOMFCNOTACPADCDPNTQASCECP 428
Db 252 CFRSKECFQCOQVNSTHYNCICAGFAKLPDDPCRCMFCNETSCPADCDPNSFCQCP 311
Qy 429 EGYILDGFICTDIDECENGFCGVCNHLNPTGFCICGPDPSALVRHIGTDCDSGKV--D 486
Db 312 EGFILDEGSICTDIDECQSGECLNECRNLPSYECICGPDPTALAGQISKDCDPIPVLED 371
Qy 487 GDSGSGGPPPS-PTPGSTLTPPVLVHSG 516
Db 372 SEDGSGEHPSSNPTVWSSTVPPSARPMHSG 402

RESULT 4
Q9HCU0 ID Q9HCU0 PRELIMINARY; PRT; 757 AA.
AC Q9HCU0;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Tumor endothelial marker 1 precursor (Endostialin protein).
GN TEM1 OR ENDOSIALIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20407466; PubMed=10947988;
RA St Croix B., Rago C., Velculescu V., Traverso G., Romans K.E.,
RA Montgomery E., Lal A., Riggins G.J., Lengauer C., Vogelstein B.,
RA Kinzler K.W.
RT "Genes expressed in human tumor endothelium.";
RL Science 289:1197-1202(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21269274; PubMed=11084048;
RA Christian S., Ahorn H., Koehler A., Eisenhaber F., Rodi H.P.,
RA Garin-Chesa P., Park J.E., Rettig W.J., Lenter M.C.;
RT "Molecular cloning and characterization of Endostialin, a C-type
lectin-like cell surface receptor of tumor Endothelium.";
RL J. Biol. Chem. 276:7408-7414(2001).
DR EMBL; AF279142; AAC00867.1;
DR EMBL; AJ295846; CAC34381.1;
DR HSSP; P07204; IZQA.
DR InterPro; IPR000152; Asx_hydroxyl.
```

```
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00008; EGF_2
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00179; EGF_CA; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00041; C-TYPE_LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 757 TUMOR ENDOTHELIAL MARKER 1.
SQ SEQUENCE 757 AA; 80858 MW; C96363EA1FD8FFA0 CRC64;

Query Match 19.9%; Score 580.5; DB 4; Length 757;
Best Local Similarity 33.4%; Pred. No. 2.5e-38;
Matches 176; Conservative 44; Mismatches 192; Indels 115; Gaps 24;

Qy 1 MLGVLVLGALALAGLG-FPAPAEPPQGGSCVHEHDCFALYPGPATFLNASQICDGLRGHL 59
Db 2 LLRLLLAWAAGPTLGQDPWAAEPR---AACGSSCTALFPRRTFLAWRACRELGGDL 58
Qy 60 MTRVSSVAADVISLLNGDGGVGRRLRWIGLQLPPGCGDKPKRLGPLRGFWVTGDNNTSY 119
Db 59 ATPRTPEARQVDSLVG--AGPASRLWIGLQAROCQOLQ--PLRGFTWTGDDQTAF 114
Qy 120 SNARLDLNGAPLGGPLCVAVSAEATVPSEPIWEEOQCEVKADGFLCEFFHPATCRPLA 179
Db 115 TNWQA-PASGGPCPAQRCAVALEAS-----GEHRWLEGSCTLAVDGYLCQFEGACPALQ 168
Qy 180 VERCAAAAVSITVGTPTFAARGADFOALPVGSSAAV--APLGLQLMCTAPPAGVQGHWA 236
Db 169 DEAGQAGPAV---YTPFHLVSTEFELWLPFGSVAAVOCQAGRGASLLCVQPEGGVG-WS 224
Qy 237 REAP---GAWDCSVENGCEHAC-NAIPGAPRCQCPAGAAALQADGRSCTASATQS-CNDL 291
Db 225 RAGPLCLGT-GCSPDNGGCEHCEVVEVDGHVSCRTGSEFRLAADGRSCDPCQAQACEQQ 283
Qy 292 CEHFCVNPDPQGSYSCHMETGYRLAADQ-HRCEVDVDDCILPSPQPCQRCVNTQGGFCH 350
Db 284 CE-----PGSPQGSYCHRCRLGFRPAEDDPRHCVDTDECCI-AGVCQCMCVNYVGSFECY 336
Qy 351 CYPNYDLVDCGEVPEVDPFCFRANCEYQCOPLNQTSYLVCAEGEAFIPHEPHRCOMFCNQ 410
Db 337 CSEGEHE-----ADGIS----- 349
Qy 411 TACPADCDP---NTQASCECEGYILDDGFICTDIDE---CENGFCGSGVCHNLPGT-- 461
Db 350 -----CSPAGAMGAQASQDLGD-ELLDDGDEDEDEAMKAFNGWT-----EMPGILW 397
Qy 462 FECICGPDPSALVRHIGTDCDSKVDGSGSGSEPPSPPTPGSTLTPP 508
Db 398 MEPTQPPDFALAYRPSFPED-----REPQIPYPEPTWPPP 432

RESULT 5
Q912V1 ID Q912V1 PRELIMINARY; PRT; 765 AA.
AC Q912V1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Endostialin.
GN TEM1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
```





[illegible]

QY	446	ENGDFCSGVCHNLPGTFCICIGPDSALVRHIGT-----DCDS-----GKVDGDSG	491
DB	393	ENNGDCSOICVNLAGSVECCQKPGFRLMKORKTCEDISECSNNGOCEQICSNOEGGYMC	442
QY	492	SCEP 495	
DB	443	SCEP 446	
RESULT 8			
ID	O88281	PRELIMINARY; PRT; 1574 AA.	
AC	O88281;		
DT	01-NOV-1998	(TrEMBLrel. 08, Created)	
DT	01-NOV-1998	(TrEMBLrel. 08, Last sequence update)	
DT	01-DEC-2001	(TrEMBLrel. 19, Last annotation update)	
DE	MEGF6.		
GN	MEGF6.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
ON	NCBI_TaxID=10116;		
OX	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;		
RX	MEDLINE=98360089; PubMed=9693030;		
RA	Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;		
RT	"Identification of high-molecular-weight proteins with multiple EGF-		
RT	like motifs by motif-trap screening.";		
RL	Genomics 51:27-34(1998).		
DR	EMBL; AB011532; BAA32462.1; -.		
DR	HSSP; P00736; IAPQ.		
DR	InterPro; IPR000152; Asx_hydroxyl.		
DR	InterPro; IPR000561; EGF-like.		
DR	InterPro; IPR001881; EGF_Ca.		
DR	Pfam; PF00008; EGF; 24.		
DR	SMART; SM00179; EGF_Ca; 4.		
DR	SMART; SM00001; EGF-like; 19.		
DR	PROSITE; PS00010; ASX_HYDROXYL; 5.		
DR	PROSITE; PS00022; EGF_1; UNKNOWN_23.		
DR	PROSITE; PS01186; EGF_2; 23.		
DR	PROSITE; PS01187; EGF_Ca; 5.		
KW	Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat		
SQ	SEQUENCE 1574 AA; 165445 MW; 2848533DBF77F67 CRC64;		
Query Match 14.2%; Score 414; DB 11; Length 1574;			
Best Local Similarity 34.1%; Pred. No. 1.3e-24;			
Matches 104; Conservative 28; Mismatches 101; Indels 72; Gaps			
QY	224	CTAPPG-AVOGHWAREAPGAWDCGVENGCEHACNAIPGAPRCQCQAGAAALQADGRSCTA	282
DB	147	CRCPFGYQLQGD-GKTCQDYDCEAHNGCGCHRCVNTPGSYLCECKPGFRLHWDGHTCL-	204
QY	283	SATQSC---NDLCEHFVCP-----NPDQGSVSCM-----309	
DB	205	-AISSTGLNGGGOHCQVLTQTHRCQCRPYQLQGEDGRRCVRRSPCAEGNGGCMHICQ	263
QY	310	-----CETGYRLAADQHRCEDVDCCILEPSPCQRCVNTQGFECHCYPNVDL-VD	359
DB	264	ELRGLAHGCGHPGYQLAADRKTCEDVDECALGLAQCAHGLCLNTQGSFKVCVCHAGVELGAD	323
QY	360	G-EG-----VEPVDPCEFRAN--CEVQCPQLNQTSLVLCVCAEGFAPIPHEPHRCQMF-	408
DB	324	GRQYRTEMEIVNSCEAGNGCGSHGSH--TSTGPLCTCPRGY-----ELDQKTCIDID	377
QY	409	---NOTACPADCPNTOA---SCPCPEGYILD-DGFTICTDDECENG-GFCSGVCNHLPGT	461
DB	378	DCANSPCCQAC-ANTPGGYECSCFAGYRLNTDGCCEVDCEASGHGCEHHCSNLAGS	436
QY	462	FEIC 466	
DB	437	FOFC 441	

RESULT 9  
ID QVSVS9 PRELIMINARY; PRT; 1394 AA.  
AC QVSVS9;  
DT 01-MAY-2000 (TREMREL. 13, Created)  
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)  
DT 01-JUN-2002 (TREMREL. 21, Last annotation update)  
DE CG7526 protein (Fragment).  
GN CG7526.

OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
ON NCBI\_TaxID=7227;  
OX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BKKELEY;  
RX MEDLINE=20196006; PubMed=10731132;

R Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Gallo R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
Wan K.H., Doyle C., Baxter E.J., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
BA Ballwey R.M., Basu A., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berhan B.P., Bhattacharya D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Douc L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fostler C., Gabrielian A.E., Gar N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Hostin N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Houston D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., MCPersonson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Neilson D.R., Neilson K.A., Nixon K., Nusser D.R., Pacle J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
DR EMBL; AE003558; AAP50338.1; .

DR HSSP; P00736; IAPQ  
DR FlyBase; FBgn0035758; Asx\_hydroxyl.  
DR InterPro; IPRO00152; Asx\_hydroxyl.  
DR InterPro; IPRO00561; EGF-like.  
DR InterPro; IPRO01881; EGF\_Ca.  
DR InterPro; IPRO03410; Hyalin.  
DR InterPro; IPRO00436; Sushi\_SCR\_CCP.  
DR InterPro; IPRO01491; Thrmomodulin.  
DR Pfam; PF00008; EGF; 11.  
DR Pfam; PF00084; sushi; 2.  
DR PRINTS; PR00907; THRMOMODULN.  
DR SMART; SM00032; CCP; 2.  
DR SMART; SM00179; EGF\_CA; 9.  
DR PROSITE; PS00001; EGF\_like; 5.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 8.

Query Match 13.2%; Score 385; DB 5; Length 1394;  
Best Local Similarity 32.1%; Pred.No. 2.4e-22;  
Matches 99; Conservative 33; Mismatches 110; Indels 66; Gaps  
SEQUENCE 1394 AA; 152269 MW; CD29380E3162F68A CRC64;

QY 224 CTAPPAGVQG-----HWAREAPGAW-----DCSVENGCEHCACNAIPG 261  
DB 563 CLCPPGYALGDNDHNIVTSLSNSFTDSSTPSAHTCLDIDECISLANGNCSHFCQNEPG 622  
QY 262 APRCOPCAPGAQAADGRSCATASATOSC---NDLCHEFCVPNPDPQSGYSWCMTCTGVRLAA 318  
DB 623 GFQCACPGLYALSDEMRTC--QQIDECLDSNGCSQLCL---NQPGFACACTGEFLTP 677  
QY 319 DOHRCEDVDCCILEPSPQRVCNTQGGFECHCYPNYDLVDGB--CVPEVPDPC---FRAN 373  
DB 678 DGFGCADIDECSDYGNCSDICINLLGTTHACACERYELAKDKLSCLD-VDECAGLLSGG 736  
QY 374 CEYQCPLNQI-TSYLCVAEGFAPITPEPHRCQMFCNQIACPA--DCDPTQTASC----- 425  
DB 737 CSHEC--INKAGTFEECGGPLY--ILNDGRR-----SCSPALVGCPPTQRSADGCAP 785  
QY 426 -ECPGTYILDGFICTDDIEC--ENGGFCGVCYNLPFTFECICGP-----DSLVRHI 476  
DB 786 IECNPGYTIGSDKKCDVIDEQKNKG-CSHRCSNTEGSFKSCPYPGYELDSQKTQCDI 844  
QY 477 GTDCDSGK 484  
DB 845 -DECDQDK 851

P87363 PRELIMINARY; PRT; 708 AA.  
AC P87363;  
DT 01-MAY-1997 (TREMREL. 03, Created)  
DT 01-MAY-1997 (TREMREL. 03, Last sequence update)  
DT 01-MAR-2002 (TREMREL. 20, Last annotation update)  
DE Fibrillin-1 (Fragment).  
GN FBNI.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelestomi;  
OC Archaeosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A..  
RX MEDLIN=20152896; PubMed=10691037;  
RA Zhou G., Price C.E., Rosencquist T.H., Gadson P.F., Godfrey M.;  
RT "Partial cloning and sequencing of chick fibrillin-1 cDNA.";  
RL In Vitro Cell. Dev. Biol. Anim. 36:19-25(2000).

DR EMBL; U88872; AAB48531.1; .  
DR HSSP; P07204; 2ADX.  
DR InterPro; IPRO00152; Asx\_hydroxyl.  
DR InterPro; IPRO00561; EGF-like.  
DR InterPro; IPRO01881; EGF\_Ca.  
DR InterPro; IPRO02212; Fibril-assoc.  
DR Pfam; PF00008; EGF; 13.  
DR Pfam; PF00683; TB; 2.  
DR PRINTS; PR00907; THRBOMODULN.  
DR SMART; SM00179; EGF\_CA; 14.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 13.  
DR PROSITE; PS01186; EGF\_2; 10.  
DR PROSITE; PS01187; EGF\_CA; 13.  
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
FT NON\_TER 708  
SQ SEQUENCE 708 AA; 76164 MW; C247271C1DF73361 CRC64;

Query Match	12.94;	Score 375;	DB 13;	Length 708;
Best Local Similarity	31.84;	Pred. No. 7.1e-22;		
Matches 112;	Conservative 26;	Mismatches 108;	Indels 106;	Gaps 20;
QY 244	DCSVENGCEHACNAIPACPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP---	300		
Db 223	ECSTMGNGCECFCTGSEGSYCSCKQGFALPDPHRTCT-----DIDE--CEDNPNIC 272			
QY 301	-----DOPGSYSCMCEGYRLADQHCEDVDCCILEPSPC-PQRCVNTGGTECHCYP 353			
Db 273	DGGQCTNIPGYRCLYDGFMASEDMKTCVDVNECDLHPNLCISGTCSTKGSFICHGDM 332			
QY 354	NYDLVDGE--CVEPVDPC--FRANCEYQCQLN-OTSILCVCAEGFA-----PIP 398			
Db 333	GYSGKKGTTGCTD-INECEIAHNCDRHAVCTNIPGSKFCSSGWIENGKICTDLDECS 391			
QY 399	HEPIRCQMFNCQTACPDPCDPTNTQAS--CECEGYILDDGFICTDIDE-----CENG- 448			
Db 392	NGTHKCSPH-----ADC-NTNGSYRCLCKEY-TGDGFTCTDLDECSNLCENGQ 442			
QY 449	-----GF-----CS-----GVCNHLPGTFECIGPDSALV 473			
Db 443	CLNAPGGYRCDCDMGLFLSLDGKACEDIDECLPNICVYGTCHNLPLGLFRCECEGVYELD 502			
QY 474	RHIG-----TDCDSK-VDGGDSGSGEPPTPGSTLTPPAVGLV 513			
Db 503	RSGGCTDVNECADPTTCISGTCVNTAGSYTCECPDP-----FELNPTRVGCV 550			
RESULT 11				
Q9WUH9	PRELIMINARY; PRT; 2906 AA.			
ID Q9WUH9	AC Q9WUH9;			
DT 01-NOV-1999	(TREMBLrel. 12, Created)			
DT 01-NOV-1999	(TREMBLrel. 12, Last sequence update)			
DT 01-MAR-2002	(TREMBLrel. 20, Last annotation update)			
OS	Fibrillin-2.			
DE	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RI	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=99350231; PubMed=10419698;			
RA	Yang Q., Ota K., Tian Y., Kumar A., Wada J., Kashiwara N., Wallner E., Kanwar Y.S.;			
RT	"Cloning of rat fibrillin-2 cDNA and its role in branching			
RL	morphogenesis of embryonic lung.";			
RL	Dev. Biol. 212:229-242(1999).			
EMBL	AFI35060; AAD34439.1; -			
HSSP	P35555; 1EMN			
DR	InterPro: IPR002086; Aldhyde_dehydr.			
DR	InterPro: IPR000152; Asx_hydroxyl.			
DR	InterPro: IPR000561; EGF-like.			
DR	InterPro: IPR001881; EGF-Ca.			
DR	InterPro: IPR001438; EGF-II.			
DR	InterPro: IPR002212; Fibril-assoc.			
DR	Pfam: PF00008; EGF; 46.			
DR	Pfam: PF00683; TB; 9.			
DR	PRINTS; PR00010; EGFBL00D.			
DR	SMART; SM00179; EGF_CA; 42.			
DR	SMART; SM00001; EGF-like; 4.			
DR	PROSITE; PS00070; ALDEHYDE_DEHYDR_CYS; UNKNOWN_1.			
DR	PROSITE; PS00010; ASX_HYDROXYL; 43.			
DR	PROSITE; PS00022; EGF_1; UNKNOWN_2.			
DR	PROSITE; PS01186; EGF-2; 36.			
DR	PROSITE; PS01187; EGF-CA; 43.			
KW	Calcium-binding; EGF-like domain.			
KW	Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.			
SEQ	SEQUENCE 2906 AA; 313371 MW; 9EE64E727044EF58 CRC64;			
Query Match	12.84;	Score 374;	DB 11;	Length 2906;
Best Local Similarity	32.94;	Pred. No. 4e-21;		
Matches 102;	Conservative 24;	Mismatches 92;	Indels 92;	Gaps 18;



ID	Q9CXD8	PRELIMINARY;	PRT;	528 AA.
AC	Q9CXDB;			
AD	01-JUN-2001 (TrEMBLrel. 17, Created)			
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)			
DE	6130401L20RIK protein.			
GN	6130401L20RIK.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=THYMUS;			
RC	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Nishikishi Y., Konno H., Adachi J., Fukuda S.,			
RA	Aizawa K., Izawa M., Fuki K., Kiyosawa H., Kondo S., Yamanaka I.,			
RA	Saito T., Okazaki Y., Gofjohori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Glasi C., King B., Kochiwa H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaudo I., Pesole G., Quackenbush J.,			
RA	Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,			
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,			
RA	Hayashizaki Y.,			
RT	"Functional annotation of a full-length mouse cDNA collection.";			
RL	Nature 409:685-690(2001).			
DR	EMBL; AK018073; BAB31061.1; -.			
DR	HSSP; P00736; IAPQ.			
DR	MGD; MGI:1922990; 6130401L20RIK.			
DR	InterPro; IPR000152; Asx_hydroxyl.			
DR	InterPro; IPR000561; EGF-like.			
DR	InterPro; IPR001881; EGF_Ca.			
DR	Pfam; PF00008; EGF; 6.			
DR	SMART; SM00181; EGF; 8.			
DR	SMART; SM00179; EGF_CA; 6.			
DR	PROSITE; PS00010; ASX_HYDROXYL; 3.			
DR	PROSITE; PS00022; EGF_1; UNKNOWN_2.			
DR	PROSITE; PS01186; EGF_2; 5.			
DR	PROSITE; PS01187; EGF_CA; 4.			
KW	EGF-like domain; Glycoprotein; Hydroxylation.			
SC	SEQUENCE 528 AA; 58217 MW; 36DE6698169328E CRC64;			
Query Match	12.3%;	Score 360;	DB 11;	Length 528;
Best Local Similarity	24.6%;	Pred. No. 8.2e-21;		
Matches 123;	Conservative 57;	Mismatches 151;	Indels 170;	Gaps
QY	86	LTWGLQLPGCGDKPLRGLRGFWVTGDNNTSYSRWARLDNGAPLGPICVAVSAABA	145	
Db	:	:	:	:
Db	12	VWAAANWPP--GSAULQLP--GMPNVCREEQTLVRLSR-----PCAA	51	
QY	146	TVPSEPIWEEQOC-----EVKADGLCFEFHPAT-----CRPLAV-----EPGAA	185	
Db	:	:	:	:
Db	52	FIDTIQFW-KQCGSPRCWGVERRIRYIIIRHYVATEHQVFRCPCGWTQWDEPGCF	110	
QY	186	AAAVSITYGTPTFAARGADFPALPVGSSAAVPLGLQMCTAPPAGVQG--HWAREAPGAWD	244	
Db	:	:	:	:
Db	111	SSLSSL--GTFHSGRECSQDTI-----RQCLCS-----QGFHGPCHOYDINE	150	
QY	245	CSVENGGC--EHACNATPGAPRCOCFAGALQADGRSC-----TASATQSC--NDL--	291	
Db	:	:	:	:
Db	151	CAVDNGGCRDRCNTI--GSYYCRCQAGOKLEEDGRGCEVDCAVVGVCQQRCLTIGT	209	
QY	292	-----CEHFCVNPDPQPGSYSCMGETGYRLAAQDHR	322	
Db	:	:	:	:



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:23:13 ; Search time 13.5 seconds  
(without alignments)  
1124.609 Million cell updates/sec

Title: US-09-509-994-1

Perfect score: 2916

Sequence: 1 MGVVLGALALAGLPAP.....PSPTGSLTPPAVLVHSG 516

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTBUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/PTBUS\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2916	100.0	575	1	US-08-261-206A-59
2	2912	99.9	575	1	US-08-312-870-1
3	2912	99.9	575	6	5466668-6
4	2904	99.6	575	1	US-08-170-290A-54
5	2842	97.5	572	6	5256770-7
6	2830	97.1	498	2	US-08-733-564-2
7	2820	96.7	497	1	US-08-312-870-3
8	2770	95.0	494	1	US-08-014-723-14
9	2770	95.0	494	1	US-08-110-011A-14
10	2768	94.9	494	1	US-08-014-723-16
11	2768	94.9	494	1	US-08-110-011A-16
12	2690	92.2	475	1	US-08-307-444A-1
13	2690	92.2	475	1	US-08-587-389-1
14	2686	92.1	475	1	US-08-307-444A-2
15	2686	92.1	475	1	US-08-587-389-2
16	2684	92.0	476	1	US-08-014-723-1
17	2684	92.0	476	1	US-08-110-011A-1
18	2682	92.0	476	1	US-08-014-723-2
19	2682	92.0	476	1	US-08-014-723-18
20	2682	92.0	476	1	US-08-110-011A-2
21	2682	92.0	476	1	US-08-110-011A-18
22	2596	89.0	456	1	US-08-307-444A-3
23	2596	89.0	456	1	US-08-587-389-3
24	2592	88.9	456	1	US-08-307-444A-4
25	2592	88.9	456	1	US-08-587-389-4
26	2543	87.2	446	1	US-08-307-444A-5
27	2543	87.2	446	1	US-08-587-389-5

28	1617	55.5	275	1	US-08-312-870-7	Sequence 7, Appl1
29	1159	39.7	215	1	US-08-312-870-5	Sequence 5, Appl1
30	685	23.5	114	2	US-08-733-564-1	Sequence 1, Appl1
31	685	23.5	115	1	US-08-312-870-9	Sequence 9, Appl1
32	580	19.9	652	2	US-08-751-305-2	Sequence 2, Appl1
33	573	19.7	492	4	US-09-724-864-39	Sequence 39, Appl1
34	358	12.3	638	2	US-08-897-443-1	Sequence 1, Appl1
35	352	12.1	58	1	US-08-261-206A-3	Sequence 3, Appl1
36	331.5	11.4	1964	4	US-09-467-997-1	Sequence 1, Appl1
37	328.5	11.3	956	2	US-08-897-443-3	Sequence 3, Appl1
38	323	11.1	1253	3	US-08-479-722B-4	Sequence 4, Appl1
39	320	11.0	1394	6	5177197-30	Patent No. 5177197
40	319.5	11.0	1833	3	US-08-479-722B-2	Sequence 2, Appl1
41	319.5	11.0	1833	5	PCT-US95-02251-18	Sequence 18, Appl1
42	315.5	10.8	443	2	US-08-833-963C-2	Sequence 2, Appl1
43	315.5	10.8	443	3	US-08-980-514-1	Sequence 1, Appl1
44	308	10.6	448	2	US-08-884-072-1	Sequence 1, Appl1
45	308	10.6	448	4	US-09-212-168-1	Sequence 1, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-261-206A-59  
; Sequence 59, Application US/08261206A  
; Patent No. 5574007

; GENERAL INFORMATION:  
; APPLICANT: Zushi, Mitichitaka  
; APPLICANT: Gomi, Komakazu  
; APPLICANT: Yamamoto, Shuji  
; APPLICANT: Suzuki, Koji  
; APPLICANT: Matsuda, Akio

; TITLE OF INVENTION: A Polypeptide Capable of Interacting  
; TITLE OF INVENTION: with Thrombin

; NUMBER OF SEQUENCES: 80

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Birch, Stewart, Kolasch & Birch

; STREET: 301 N. Washington St.

; CITY: Falls Church

; STATE: Virginia

; COUNTRY: USA

; ZIP: 22046-0747

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patencin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/261.206A

; FILING DATE:

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/740,492

; FILING DATE: 03-AUG-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: Svensson, Leonard R.

; REGISTRATION NUMBER: 30330

; REFERENCE/DOCKET NUMBER: 216-275P

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-241-1300

; TELEFAX: 703-241-2848

; TELEX: 248345

; INFORMATION FOR SEQ ID NO: 59:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 575 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; FEATURE:

NO HETEROLOGOUS SEQUENCES

```

; NAME/KEY: Protein
; LOCATION: 1..575
; OTHER INFORMATION: /label= protein
; OTHER INFORMATION: /note= "human thrombomodulin"
US-08-261-206A-59

Query Match      100.0%; Score 2916; DB 1; Length 575;
Best Local Similarity 100.0%; Pred. No. 1.2e-198;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVLGALALAGLGFPAEPQPGSGQCVHEHDCFALYPGPATFELNASQICDGLRGHLM 60
DB 1 MLGVLVLGALALAGLGFPAEPQPGSGQCVHEHDCFALYPGPATFELNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQLPPGCGDKPRGLGRGFWMTGDNNTSYS 120
DB 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQLPPGCGDKPRGLGRGFWMTGDNNTSYS 120
QY 121 RWARDLNGAPLCGLPCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEFFHFPATCRPLAY 180
DB 121 RWARDLNGAPLCGLPCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEFFHFPATCRPLAY 180
QY 181 EPGAAAAYSIITYGTFFAARGADFOALPYGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAAYSIITYGTFFAARGADFOALPYGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGGECHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAWDCSVENGGECHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCETGYRLAADOHRCEDVDDCILEPSPQPCQVNTQGGFECHECPNYDLVDG 360
DB 301 DQPGSYSCMCETGYRLAADOHRCEDVDDCILEPSPQPCQVNTQGGFECHECPNYDLVDG 360
QY 361 ECVEPVDPFCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCQMFCTACPADCDPN 420
DB 361 ECVEPVDPFCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCQMFCTACPADCDPN 420
QY 421 TQASCEPCEGYILLDDGFICTDIDECENGFGCSGVCHNLPGTFECICGPDSSALVRHIGTDC 480
DB 421 TQASCEPCEGYILLDDGFICTDIDECENGFGCSGVCHNLPGTFECICGPDSSALVRHIGTDC 480
QY 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 2
US-08-312-870-1
; Sequence 1, Application US/08312870
; Patent No. 5639625
; GENERAL INFORMATION:
; APPLICANT: Carson, Craig W.
; APPLICANT: Esmen, Charles T.
; TITLE OF INVENTION: Method for Detecting Antibodies to
; TITLE OF INVENTION: Thrombomodulin in Patients
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richards, Medlock & Andrews
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: Texas
; COUNTRY: US
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,870
; FILING DATE:
; CLASSIFICATION: 435
```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Hansen, Eugenia S.
; REGISTRATION NUMBER: 31,966
; REFERENCE/DOCKET NUMBER: OMRF B35150
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-939-4500
; TELEFAX: 214-939-4600
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 19..575
US-08-312-870-1

Query Match      99.9%; Score 2912; DB 1; Length 575;
Best Local Similarity 99.8%; Pred. No. 2.3e-198;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGVLVLGALALAGLGFPAEPQPGSGQCVHEHDCFALYPGPATFELNASQICDGLRGHLM 60
DB 1 MLGVLVLGALALAGLGFPAEPQPGSGQCVHEHDCFALYPGPATFELNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQLPPGCGDKPRGLGRGFWMTGDNNTSYS 120
DB 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQLPPGCGDKPRGLGRGFWMTGDNNTSYS 120
QY 121 RWARDLNGAPLCGLPCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEFFHFPATCRPLAY 180
DB 121 RWARDLNGAPLCGLPCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEFFHFPATCRPLAY 180
QY 181 EPGAAAAYSIITYGTFFAARGADFOALPYGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAAYSIITYGTFFAARGADFOALPYGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGGECHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAWDCSVENGGECHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCETGYRLAADOHRCEDVDDCILEPSPQPCQVNTQGGFECHECPNYDLVDG 360
DB 301 DQPGSYSCMCETGYRLAADOHRCEDVDDCILEPSPQPCQVNTQGGFECHECPNYDLVDG 360
QY 361 ECVEPVDPFCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCQMFCTACPADCDPN 420
DB 361 ECVEPVDPFCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCQMFCTACPADCDPN 420
QY 421 TQASCEPCEGYILLDDGFICTDIDECENGFGCSGVCHNLPGTFECICGPDSSALVRHIGTDC 480
DB 421 TQASCEPCEGYILLDDGFICTDIDECENGFGCSGVCHNLPGTFECICGPDSSALVRHIGTDC 480
QY 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 3
5466668-6
; Patent No. 5466668
; APPLICANT: GLASER, CHARLES B.;MORSER, MICHAEL J.;LIGHT,
; DAVID R.
; TITLE OF INVENTION: SUPERIOR THROMBOMODULIN ANALOGS FOR
; PHARMACEUTICAL USE
; NUMBER OF SEQUENCES: 57
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/155,346
; FILING DATE: 22-NOV-1993
```



```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 568,456
; FILING DATE: 15-AUG-1990
; APPLICATION NUMBER: 506,325
; FILING DATE: 09-APR-1990
; APPLICATION NUMBER: 406,941
; FILING DATE: 13-SEP-1989
; APPLICATION NUMBER: 345,374
; FILING DATE: 28-APR-1989
; SEQ ID NO:6:
; LENGTH: 575
5466668-6

Query Match          99.9%; Score 2912; DB 6; Length 575;
Best Local Similarity 99.8%; Pred. No. 2.3e-198;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1  MGVLVLGALAGLGGFPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRHLM 60
DB 1  MGVLVLGALAGLGGFPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRHLM 60

QY 61 TVRSSVAADVITSLINGDGGVRRRLWTGLPPCGGDPKRLGRLGFWTGDNTSYS 120
DB 61 TVRSSVAADVITSLINGDGGVRRRLWTGLPPCGGDPKRLGRLGFWTGDNTSYS 120

QY 121 RWARDLNGAPLCGPLCVAVSAAEATVPSEPIWEQQCEVKADGFLCFEFHPATCRPLAV 180
DB 121 RWARDLNGAPLCGPLCVAVSAAEATVPSEPIWEQQCEVKADGFLCFEFHPATCRPLAV 180

QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAP 240

QY 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEFCVNP 300
DB 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEFCVNP 300

QY 301 DPGSYSCMCTGYRLAADHRCEDVDDCILEPSPCORCVNTQGGFCHCYPNYDLVDG 360
DB 301 DPGSYSCMCTGYRLAADHRCEDVDDCILEPSPCORCVNTQGGFCHCYPNYDLVDG 360

QY 361 ECVEPVPDPCFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
DB 361 ECVEPVPDPCFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420

QY 421 TQASCECPGYILDDGFICTDIDECENGCGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDC 480
DB 421 TQASCECPGYILDDGFICTDIDECENGCGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDC 480

QY 481 DSGKVDGDSGSGEPPSPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGDSGSGEPPSPPTPGSTLTTPPAVGLVHSG 516
```

## RESULT 4

```
US-08-170-290A-54
; Sequence 54; Application US/08170290A
; Patent No. 5702931
; GENERAL INFORMATION:
; APPLICANT: Andrews, William H.
; APPLICANT: Morser, Michael J.
; APPLICANT: Zieglender, Laura R.
; TITLE OF INVENTION: No. 5702931el Mutagenesis Methods and
; TITLE OF INVENTION: Compositions
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James M. Heslin
; STREET: 379 Lytton Ave.
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/170,290A
; FILING DATE: 28-DEC-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/05573
; FILING DATE: 01-JUL-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/724,237
; FILING DATE: 01-JUL-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Heslin, James M.
; REGISTRATION NUMBER: 29,541
; REFERENCE/DOCKET NUMBER: 11973-58-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 575 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-170-290A-54

Query Match          99.6%; Score 2904; DB 1; Length 575;
Best Local Similarity 99.6%; Pred. No. 8.4e-198;
Matches 514; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1  MGVLVLGALAGLGGFPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRHLM 60
DB 1  MGVLVLGALAGLGGFPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRHLM 60

QY 61 TVRSSVAADVITSLINGDGGVRRRLWTGLPPCGGDPKRLGRLGFWTGDNTSYS 120
DB 61 TVRSSVAADVITSLINGDGGVRRRLWTGLPPCGGDPKRLGRLGFWTGDNTSYS 120

QY 121 RWARDLNGAPLCGPLCVAVSAAEATVPSEPIWEQQCEVKADGFLCFEFHPATCRPLAV 180
DB 121 RWARDLNGAPLCGPLCVAVSAAEATVPSEPIWEQQCEVKADGFLCFEFHPATCRPLAV 180

QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHWAREAP 240

QY 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEFCVNP 300
DB 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEFCVNP 300

QY 301 DPGSYSCMCTGYRLAADHRCEDVDDCILEPSPCORCVNTQGGFCHCYPNYDLVDG 360
DB 301 DPGSYSCMCTGYRLAADHRCEDVDDCILEPSPCORCVNTQGGFCHCYPNYDLVDG 360

QY 361 ECVEPVPDPCFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
DB 361 ECVEPVPDPCFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420

QY 421 TQASCECPGYILDDGFICTDIDECENGCGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDC 480
DB 421 TQASCECPGYILDDGFICTDIDECENGCGFCGVCVCHNLPGTFECICGPDLSALVRHIGTDC 480

QY 481 DSGKVDGDSGSGEPPSPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGDSGSGEPPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 5
```

```
5256770-7
; Patent No. 5256770
; APPLICANT: GLASER, CHARLES B.; MORSE, MICHAEL J.; LIGHT, DAVID R.
; TITLE OF INVENTION: OXIDATION RESISTANT THROMBOMODULIN ANALOGS
; NUMBER OF SEQUENCES: 48
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/506,325
; FILING DATE: 09-APR-1990
; SEQ ID NO: 7:
; LENGTH: 572
5256770-7

Query Match          97.5%; Score 2842.5; DB 6; Length 572;
Best Local Similarity 98.4%; Pred. No. 1.8e-193;
Matches 508; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 1 MLGVLVGLALAGLGFPAEPAPQPGSQCVHDCFALYGPATFLNASQICDGLRGLHLM 60
DB 1 MLGVLVGLALAGLGFPAEPAPQPGSQCVHDCFALYGPATFLNASQICDGLRGLHLM 60
QY 61 TVRSSVAADVISLLNGDGVGRRRLWIGLQLPPGCGDKRRLGRLGFWQVTTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGVG--RRLWIGLQLPPGCGDKRRLGRLGFWQVTTGDNNTSYS 119
QY 121 RWARLDLNGAPLCPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFHFPATCRPLAV 180
DB 121 RWARLDLNGAPLCPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFHFPATCRPLAV 179
QY 181 EPGAAAVSITTYGTPFAAGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
DB 180 EPGAAAVSITTYGTPFAAGADFOALPVGSSAAVAPLGLQMLCTA--GNVQGHWAREAP 237
QY 241 GAWDCSVENGCGCHEACNAIFCAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFVCPNP 300
DB 238 GAWDCSVENGCGCHEACNAIFCAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFVCPNP 297
QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYNYDLVDG 360
DB 298 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYNYDLVDG 357
QY 361 ECVEPVPDPCFRANCEYOCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420
DB 358 ECVEPVPDPCFRANCEYOCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 417
QY 421 TQASCECEGYILLDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 480
DB 418 TQASCECEGYILLDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDC 477
QY 481 DSGKVDGSDSGSGPPSPPTPGSTLTTPPAVGLVHSG 516
DB 478 DSGKVDGSDSGSGPPSPPTPGSTLTTPPAVGLVHSG 513

RESULT 6
US-08-733-564-2
; Sequence 2, Application US/08733564
; Patent No. 5916874
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, Kenji
; APPLICANT: MOCHIDA, Satoshi
; TITLE OF INVENTION: METHOD FOR TREATING LIVER INJURY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/733,564
; FILING DATE: 18 OCTOBER 1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, Leonard R.
; REGISTRATION NUMBER: 30,330
; REFERENCE/DOCKET NUMBER: 0216-0362P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 498 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-08-733-564-2

Query Match          97.1%; Score 2830; DB 2; Length 498;
Best Local Similarity 100.0%; Pred. No. 1.2e-192;
Matches 498; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 APAPQPGSQCVHDCFALYGPATFLNASQICDGLRGLHLMTVRSSVAADVISLLNGD 78
DB 1 APAPQPGSQCVHDCFALYGPATFLNASQICDGLRGLHLMTVRSSVAADVISLLNGD 60
QY 79 GVGRRRLWIGLQLPPGCGDKRRLGRLGFWQVTTGDNNTSYSRWARLDLNGAPLCPLCV 138
DB 61 GVGRRRLWIGLQLPPGCGDKRRLGRLGFWQVTTGDNNTSYSRWARLDLNGAPLCPLCV 120
QY 139 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFHFPATCRPLAVEPGAAAAVSIYGTTPFA 198
DB 121 AVSAAEATVPSEPIWEEQCEVKADGFLCEHFHFPATCRPLAVEPGAAAAVSIYGTTPFA 180
QY 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCGHEACNA 258
DB 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCGHEACNA 240
QY 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFVCPNPDPGSGYSCMCTGYRLAA 318
DB 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFVCPNPDPGSGYSCMCTGYRLAA 300
QY 319 DQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYNYDLVDGECVEPVPDPCFRANCEYOC 378
DB 301 DQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYNYDLVDGECVEPVPDPCFRANCEYOC 360
QY 379 QPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDNTQASCECPGEGYILLDDGFI 438
DB 361 QPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDNTQASCECPGEGYILLDDGFI 420
QY 439 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDQSGKVDGSDSGSGPPSP 498
DB 421 CTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRHIGTDCDQSGKVDGSDSGSGPPSP 480
QY 499 PTPGSTLTTPPAVGLVHSG 516
DB 481 PTPGSTLTTPPAVGLVHSG 498

RESULT 7
US-08-312-870-3
; Sequence 3, Application US/08312870
; Patent No. 5639625
; GENERAL INFORMATION:
; APPLICANT: Carson, Craig W.
; APPLICANT: Esmon, Charles T.
; TITLE OF INVENTION: Method for Detecting Antibodies to Thrombomodulin in Patients
; NUMBER OF SEQUENCES: 11
```

CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Richards, Medlock & Andrews  
 STREET: 1201 Elm Street, Suite 4500  
 CITY: Dallas  
 STATE: Texas  
 COUNTRY: US  
 ZIP: 75270-2197  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/312,870  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Hansen, Eugenia S.  
 REGISTRATION NUMBER: 31,966  
 REFERENCE/DOCKET NUMBER: OMRF B35150  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 214-939-4500  
 TELEFAX: 214-939-4600  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 497 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-312-870-3

Query Match 96.7%; Score 2820; DB 1; Length 497;  
 Best Local Similarity 99.8%; Pred. No. 6.2e-192;  
 Matches 496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 19 APAEPQGGSCVCHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 78  
 DB 1 APAEPQGGSCVCHDCFALYPGPATFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 60  
 QY 79 GGVRRRLWIGLQPLPGCGDKRGLRGFQWVGTGDNNTSYSRWARLDLNGAPLCGPCLV 138  
 DB 61 GGVRRRLWIGLQPLPGCGDKRGLRGFQWVGTGDNNTSYSRWARLDLNGAPLCGPCLV 120  
 QY 139 AVSAAEATVPSEPIWEEOCEVKADGFLCEHFPATCRPLAVEPGAAAAAIVTGTPEFA 198  
 DB 121 AVSAAEATVPSEPIWEEOCEVKADGFLCEHFPATCRPLAVEPGAAAAAIVTGTPEFA 180  
 QY 199 ARGADFQALPVGSSAAVAPLGQLMCTAPPGAVQGHWAREAPGAWDCSVENGCGCEHACNA 258  
 DB 181 ARGADFQALPVGSSAAVAPLGQLMCTAPPGAVQGHWAREAPGAWDCSVENGCGCEHACNA 240  
 QY 259 IPGAPRCQCPAGAAQAADGRSCTASATOSCNLCEHFCVFPNPDQPGSYSCMCETGYRLAA 318  
 DB 241 IPGAPRCQCPAGAAQAADGRSCTASATOSCNLCEHFCVFPNPDQPGSYSCMCETGYRLAA 300  
 QY 319 DQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYNYDLVDGECVEPVDPCFRANCEYQC 378  
 DB 301 DQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYNYDLVDGECVEPVDPCFRANCEYQC 360  
 QY 379 QPLNQTSLYLCVCAEGFAPLPHBPHRCQMFNCNOTACPADCPMTQASCEPGEYILDDGFI 438  
 DB 361 QPLNQTSLYLCVCAEGFAPLPHBPHRCQMFNCNOTACPADCPMTQASCEPGEYILDDGFI 420  
 QY 439 CTFDIDECNGGFCGVCNHLPTFCICGPDPSALYRHGTTDCDSKVDGSDSGSGEPPPS 498  
 DB 421 CTFDIDECNGGFCGVCNHLPTFCICGPDPSALYRHGTTDCDSKVDGSDSGSGEPPPS 480  
 QY 499 PTPGSTLTTPPAVGLVHS 515  
 DB 481 PTPGSTLTTPPAVGLVHS 497

RESULT 8

US-08-014-723-14  
 Sequence 14, Application US/08014723  
 Patent No. 5273962  
 GENERAL INFORMATION:  
 APPLICANT: DOI, Takeshi  
 APPLICANT: Iwasaki, Akio  
 APPLICANT: Saino, Yushi  
 APPLICANT: Kimura, Shigeru  
 APPLICANT: Ohkuchi, Masao  
 TITLE OF INVENTION: Thrombin-Binding Substance and Process  
 TITLE OF INVENTION: For Preparing the Same  
 NUMBER OF SEQUENCES: 18  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
 ADDRESSEE: P.C.  
 STREET: 1755 Jefferson Davis Highway, Fourth Floor  
 CITY: Arlington  
 STATE: Virginia  
 ZIP: 22202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/014,723  
 FILING DATE: 19930208  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Oblon, No. 5273962man F.  
 REGISTRATION NUMBER: 24,618  
 REFERENCE/DOCKET NUMBER: 80-071-0 CIP  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703)413-3000  
 TELEFAX: (703)413-2220  
 TELEX: 248855 OPAT UR  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 494 amino acids  
 TYPE: AMINO ACID  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-014-723-14

Query Match 95.0%; Score 2770; DB 1; Length 494;  
 Best Local Similarity 99.4%; Pred. No. 2.1e-188;  
 Matches 491; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 MLGVLVLGALALAGLGFPAEPQPGSQCVCHDCFALYPGPATFLNASQICDGLRGLH 60  
 DB 1 MLGVLVLGALALAGLGFPAEPQPGSQCVCHDCFALYPGPATFLNASQICDGLRGLH 60  
 QY 61 TVRSSVAADVLSLLNGDGGVGRRLWTGLQPLPGCGDKRGLRGFQWVGTGDNNTSY 120  
 DB 61 TVRSSVAADVLSLLNGDGGVGRRLWTGLQPLPGCGDKRGLRGFQWVGTGDNNTSY 120  
 QY 121 RWARDLNGAPLCGPLCVAVSAAEATVPSEPIWEEOCEVKADGFLCEHFPATCRPLAV 180  
 DB 121 RWARDLNGAPLCGPLCVAVSAAEATVPSEPIWEEOCEVKADGFLCEHFPATCRPLAV 180  
 QY 181 EPGAAAAVSYTGTPTFAARGADFOALPVGSSAAVAPLGQLMCTAPPGAVQGHWAREAP 240  
 DB 181 EPGAAAAVSYTGTPTFAARGADFOALPVGSSAAVAPLGQLMCTAPPGAVQGHWAREAP 240  
 QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQAADGRSCTASATOSCNLCEHFCV 300  
 DB 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQAADGRSCTASATOSCNLCEHFCV 300  
 QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYNYDLVDG 360  
 DB 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYNYDLVDG 360  
 QY 361 ECVPEVDPFCFRANCEYQCPLNQTSLYLCVCAEGFAPLPHBPHRCQMFNCNOTACPADCPDN 420

Db 361 ECVPEVDPFRANCEYQCQLNQTSLVCVAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
QY 421 TQASCEPGEYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRRHIGTDC 480  
Db 421 TQASCEPGEYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRRHIGTDC 480  
QY 481 DSGKVDGDSGSGE 494  
Db 481 DSGKVDGDSGSGE 494

RESULT 9  
US-08-110-011A-14  
; Sequence 14, Application US/08110011A  
; Patent No. 5354664  
; GENERAL INFORMATION:  
; APPLICANT: Doi, Takeshi  
; APPLICANT: Iwasaki, Akio  
; APPLICANT: Saino, Yushi  
; APPLICANT: Kimura, Shigeru  
; APPLICANT: Ohkuchi, Masao  
; TITLE OF INVENTION: Thrombin-Binding Substance and Process  
; TITLE OF INVENTION: For Preparing the Same  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/110,011A  
; FILING DATE: 23-AUG-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5354664man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 80-073-0 DIV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)413-3000  
; TELEFAX: (703)413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 494 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-110-011A-14

Query Match 95.0%; Score 2770; DB 1; Length 494;  
Best Local Similarity 99.4%; Pred. No. 2.1e-188;  
Matches 491; Conservative 0; Mismatches 3; Indels 0; Caps 0;  
QY 1 MLGVLVLAGALAGLFPAPAEPPQGSQCVEHDCFALYFGPATFLNASQICDGLRHL 60  
Db 1 MLGVLVLAGALAGLFPAPAEPPQGSQCVEHDCFALYFGPATFLNASQICDGLRHL 60  
QY 61 TVRSSVAADVLSILLNGDGVGRRLRWIGLQLPFGCDPRRLGRLGFWVTGDNNTSYS 120  
Db 61 TVRSSVAADVLSILLNGDGVGRRLRWIGLQLPFGCDPRRLGRLGFWVTGDNNTSYS 120  
QY 121 RWARLDLNGAPLCGLCVAVSAEAATVPSEPIWEEQCEVKADGFLCEHFHFPATCRPLAV 180  
Db 121 RWARLDLNGAPLCGLCVAVSAEAATVPSEPIWEEQCEVKADGFLCEHFHFPATCRPLAV 180

QY 181 EPGAAAAVSIYGTTPFAARGADFOALPYGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
Db 181 EPGAAAAVSIYGTTPFAARGADFOALPYGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
QY 241 GAWDCSVENGCGGHACNAIPGAPRCOCPCAGAAQAQDGRSCTASATOSCNLDCHEHFCVNP 300  
Db 241 GAWDCSVENGCGGHACNAIPGAPRCOCPCAGAAQAQDGRSCTASATOSCNLDCHEHFCVNP 300  
QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDDCILEPSPQRCVNTQGGFECHECHYNYDLVDG 360  
Db 301 DQPGSYSCMCETGYRLAADQHRCEVDVDDCILEPSPQRCVNTQGGFECHECHYNYDLVDG 360  
QY 361 ECVPEVDPFRANCEYQCQLNQTSLVCVAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
Db 361 ECVPEVDPFRANCEYQCQLNQTSLVCVAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
QY 421 TQASCEPGEYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRRHIGTDC 480  
Db 421 TQASCEPGEYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVRRHIGTDC 480  
QY 481 DSGKVDGDSGSGE 494  
Db 481 DSGKVDGDSGSGE 494  
RESULT 10  
US-08-014-723-16  
; Sequence 16, Application US/08014723  
; Patent No. 5273962  
; GENERAL INFORMATION:  
; APPLICANT: Doi, Takeshi  
; APPLICANT: Iwasaki, Akio  
; APPLICANT: Saino, Yushi  
; APPLICANT: Kimura, Shigeru  
; APPLICANT: Ohkuchi, Masao  
; TITLE OF INVENTION: Thrombin-Binding Substance and Process  
; TITLE OF INVENTION: For Preparing the Same  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/014,723  
; FILING DATE: 19930208  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5273962man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 80-071-0 CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)413-3000  
; TELEFAX: (703)413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 494 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-014-723-16

Query Match 94.9%; Score 2768; DB 1; Length 494;  
Best Local Similarity 99.2%; Pred. No. 2.9e-188;  
Matches 490; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MLGVLVGLALAGLGFPAEAPQPGSGQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
Db 1 MLGVLVGLALAGLGFPAEAPQPGSGQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFWQWVTGDNNTSYS 120  
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFWQWVTGDNNTSYS 120  
QY 121 RWARLDLNGAPLCGLPCLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
Db 121 RWARLDLNGAPLCGLPCLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQMLCTAPPAGAVQGHWAREAP 240  
Db 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQMLCTAPPAGAVQGHWAREAP 240  
QY 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
Db 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DPGSYSCMCEGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHCYPNYDLVDG 360  
Db 301 DPGSYSCMCEGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHCYPNYDLVDG 360  
QY 361 ECVEPVDPFCFRANCEYQCPNLQNTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCPN 420  
Db 361 ECVEPVDPFCFRANCEYQCPNLQNTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCPN 420  
QY 421 TOASCCEPGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVHRHIGTDC 480  
Db 421 TOASCCEPGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVHRHIGTDC 480  
QY 481 DSGKVDGDSGSGE 494  
Db 481 DSGKVDDEASGSGD 494

RESULT 11  
US-08-110-011A-16  
; Sequence 16, Application US/08110011A  
; Patent No. 5354664  
; GENERAL INFORMATION:  
; APPLICANT: Doi, Takeshi  
; APPLICANT: Iwasaki, Akio  
; APPLICANT: Saino, Yushi  
; APPLICANT: Kimura, Shigeru  
; APPLICANT: Okuchi, Masao  
; TITLE OF INVENTION: Thrombin-Binding Substance and Process  
; TITLE OF INVENTION: For Preparing the Same  
; NUMBER OF SEQUENCES: 18  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OBLON, SPIVAK, McLELLAND, MAIER & NEUSTADT,  
; ADDRESSER: P.C.  
; STREET: 1755 Jefferson Davis Highway, Fourth Floor  
; CITY: Arlington  
; STATE: Virginia  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/110,011A  
; FILING DATE: 23-AUG-1993  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Oblon, No. 5354664man F.  
; REGISTRATION NUMBER: 24,618  
; REFERENCE/DOCKET NUMBER: 80-073-0 DIV  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)413-3000

; TELEFAX: (703)413-2220  
; TELEX: 248855 OPAT UR  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 494 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-110-011A-16  
  
Query Match 94.9%; Score 2768; DB 1; Length 494;  
Best Local Similarity 99.2%; Pred. No. 2.9e-188;  
Matches 490; Conservative 1; Mismatches 3; Indels 0; Gaps 0:  
  
QY 1 MLGVLVGLALAGLGFPAEAPQPGSGQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
Db 1 MLGVLVGLALAGLGFPAEAPQPGSGQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFWQWVTGDNNTSYS 120  
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFWQWVTGDNNTSYS 120  
QY 121 RWARLDLNGAPLCGLPCLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
Db 121 RWARLDLNGAPLCGLPCLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQMLCTAPPAGAVQGHWAREAP 240  
Db 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQMLCTAPPAGAVQGHWAREAP 240  
QY 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
Db 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DPGSYSCMCEGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHCYPNYDLVDG 360  
Db 301 DPGSYSCMCEGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECCHCYPNYDLVDG 360  
QY 361 ECVEPVDPFCFRANCEYQCPNLQNTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCPN 420  
Db 361 ECVEPVDPFCFRANCEYQCPNLQNTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCPN 420  
QY 421 TOASCCEPGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVHRHIGTDC 480  
Db 421 TOASCCEPGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALVHRHIGTDC 480  
QY 481 DSGKVDGDSGSGE 494  
Db 481 DSGKVDDEASGSGD 494

RESULT 12  
US-08-307-444A-1  
; Sequence 1, Application US/08307444A  
; Patent No. 5516559  
; GENERAL INFORMATION:  
; APPLICANT: NII, ATSUSHI  
; APPLICANT: MORISHITA, HIDEAKI  
; APPLICANT: UEMURA, AKIO  
; APPLICANT: MOCHIDA, EI  
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OLIFF & BERRIDGE  
; STREET: P.O. BOX 19928  
; CITY: ALEXANDRIA  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22320  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/307,444A  
; FILING DATE: 19-SEP-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/835,436  
; FILING DATE: 26-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OLIFF, JAMES A.  
; REGISTRATION NUMBER: 27,075  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6400  
; TELEFAX: (703) 836-2787  
; TELEX: 90-1799 PTO ALEX  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 475 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; US-08-307-444A-1

Query Match 92.2%; Score 2690; DB 1; Length 475;  
Best Local Similarity 100.0%; Pred. No. 9e-183;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MLGVLVLGALALAGLGFAPAPPPGSGQCVHEHDFALYPGPATFLNASQICDGLRGHLM 60  
Db 1 MLGVLVLGALALAGLGFAPAPPPGSGQCVHEHDFALYPGPATFLNASQICDGLRGHLM 60  
  
QY 61 TVRSSVAADVISLLNGDGGVGRRLWIGLQLPPGCGDPKRLGRLGFQWVTGDNNTSYS 120  
Db 61 TVRSSVAADVISLLNGDGGVGRRLWIGLQLPPGCGDPKRLGRLGFQWVTGDNNTSYS 120  
  
QY 121 RWARDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAY 180  
Db 121 RWARDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAY 180  
  
QY 181 EPGAAAAVSIYTGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240  
Db 181 EPGAAAAVSIYTGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240  
  
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
  
QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
Db 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
  
QY 361 ECVEPVPDCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420  
Db 361 ECVEPVPDCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420  
  
QY 421 TQASCECPGYLLDDGFICTDIDECENGFGCSGVCHNLPGTPEICGPDPSALVRH 475  
Db 421 TQASCECPGYLLDDGFICTDIDECENGFGCSGVCHNLPGTPEICGPDPSALVRH 475

## RESULT 13

US-08-587-389-1  
; Sequence 1, Application US/08587389  
; Patent No. 5695964  
; GENERAL INFORMATION:  
; APPLICANT: NII, ATSUSHI  
; APPLICANT: MORISHITA, HIDEAKI  
; APPLICANT: UEMURA, AKIO  
; APPLICANT: MOCHIDA, EI

; TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT  
; TITLE OF INVENTION: PRODUCTION THEREOF, AND THERAPEUTIC AGENT (AS AMENDED)  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OLIFF & BERRIDGE  
; STREET: P.O. BOX 19928  
; CITY: ALEXANDRIA  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22320  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/587,389  
; FILING DATE: 17-JAN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/307,444  
; FILING DATE: 19-SEP-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OLIFF, JAMES A.  
; REGISTRATION NUMBER: 27,075  
; REFERENCE/DOCKET NUMBER: JAO 27706  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6400  
; TELEFAX: (703) 836-2787  
; TELEX: 90-1799 PTO ALEX  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 475 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; US-08-587-389-1

Query Match 92.2%; Score 2690; DB 1; Length 475;  
Best Local Similarity 100.0%; Pred. No. 9e-183;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MLGVLVLGALALAGLGFAPAPPPGSGQCVHEHDFALYPGPATFLNASQICDGLRGHLM 60  
Db 1 MLGVLVLGALALAGLGFAPAPPPGSGQCVHEHDFALYPGPATFLNASQICDGLRGHLM 60  
  
QY 61 TVRSSVAADVISLLNGDGGVGRRLWIGLQLPPGCGDPKRLGRLGFQWVTGDNNTSYS 120  
Db 61 TVRSSVAADVISLLNGDGGVGRRLWIGLQLPPGCGDPKRLGRLGFQWVTGDNNTSYS 120  
  
QY 121 RWARDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAY 180  
Db 121 RWARDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQCEVKADGFLCFEHPFATCRPLAY 180  
  
QY 181 EPGAAAAVSIYTGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240  
Db 181 EPGAAAAVSIYTGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240  
  
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
  
QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
Db 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
  
QY 361 ECVEPVPDCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420  
Db 361 ECVEPVPDCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420

QY 421 TQASCEPEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFTCEICGPDPSALVRH 475  
|||||  
Db 421 TQASCEPEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFTCEICGPDPSALVRH 475  
|||||  
RESULT 14  
US-08-307-444A-2  
; Sequence 2, Application US/08307444A  
; Patent No. 5516659  
; GENERAL INFORMATION:  
; APPLICANT: NII, ATSUSHI  
; APPLICANT: MORISHITA, HIDEAKI  
; APPLICANT: UEMURA, AKIO  
; APPLICANT: MOCHIDA, EI  
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OLIFF & BERRIDGE  
; STREET: P.O. BOX 19928  
; CITY: ALEXANDRIA  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22320  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/307,444A  
; FILING DATE: 19-SEP-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/835,436  
; FILING DATE: 26-FEB-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OLIFF, JAMES A.  
; REGISTRATION NUMBER: 27,075  
; REFERENCE/DOCKET NUMBER: JAO 27706  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6400  
; TELEFAX: (703) 836-2787  
; TELEX: 90-1799 PTO ALEX  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 475 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-307-444A-2  
Query Match 92.1%; Score 2686; DB 1; Length 475;  
Best Local Similarity 99.8%; Pred. No. 1.7e-182;  
Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MLGVVLGALALAGLGPAPAEPPQSGQCVHEHDCFALYPGPATFLNASQICDGLRGLHLM 60  
|||||  
Db 1 MLGVVLGALALAGLGPAPAEPPQSGQCVHEHDCFALYPGPATFLNASQICDGLRGLHLM 60  
|||||  
QY 61 TVRSSVAADVLSLLNDGGVGRRLRIGLQLPPCGDPRKRLGPIRGFWQWTGDNNTSYS 120  
|||||  
Db 61 TVRSSVAADVLSLLNDGGVGRRLRIGLQLPPCGDPRKRLGPIRGFWQWTGDNNTSYS 120  
|||||  
QY 121 RWARLDLNGAPLCGPLCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEFFHPATCRPLAY 180  
|||||  
Db 121 RWARLDLNGAPLCGPLCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEFFHPATCRPLAY 180  
|||||  
QY 181 EPGAAAASVITYGTPFAARGADPQALPVGSSAAVAPLGLQLMCTAPPVAGVGHWAREAP 240  
|||||  
Db 181 EPGAAAASVITYGTPFAARGADPQALPVGSSAAVAPLGLQLMCTAPPVAGVGHWAREAP 240  
|||||

QY 241 GAWDCSVENGCGEACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVPNP 300  
|||||  
Db 241 GAWDCSVENGCGEACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVPNP 300  
|||||  
QY 301 DOPGYSYSCMCTGYRLAADOHRCEVDVDDCILEPSPCPQRCVNTQGGFECGCHYPNYDLVDG 360  
|||||  
Db 301 DOPGYSYSCMCTGYRLAADOHRCEVDVDDCILEPSPCPQRCVNTQGGFECGCHYPNYDLVDG 360  
|||||  
QY 361 ECVEPVDFCFRANCEYQCOPLNQTSLYCVCAEGFAPIPHEPHRCQMFNQTACPADCDPN 420  
|||||  
Db 361 ECVEPVDFCFRANCEYQCOPLNQTSLYCVCAEGFAPIPHEPHRCQMFNQTACPADCDPN 420  
|||||  
QY 421 TQASCEPEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFTCEICGPDPSALVRH 475  
|||||  
Db 421 TQASCEPEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFTCEICGPDPSALVRH 475  
|||||  
RESULT 15  
US-08-587-389-2  
; Sequence 2, Application US/08587389  
; Patent No. 5695964  
; GENERAL INFORMATION:  
; APPLICANT: NII, ATSUSHI  
; APPLICANT: MORISHITA, HIDEAKI  
; APPLICANT: UEMURA, AKIO  
; APPLICANT: MOCHIDA, EI  
; TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: OLIFF & BERRIDGE  
; STREET: P.O. BOX 19928  
; CITY: ALEXANDRIA  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22320  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/587,389  
; FILING DATE: 17-JAN-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/307,444  
; FILING DATE: 19-SEP-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OLIFF, JAMES A.  
; REGISTRATION NUMBER: 27,075  
; REFERENCE/DOCKET NUMBER: JAO 27706  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6400  
; TELEFAX: (703) 836-2787  
; TELEX: 90-1799 PTO ALEX  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 475 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
US-08-587-389-2  
Query Match 92.1%; Score 2686; DB 1; Length 475;  
Best Local Similarity 99.8%; Pred. No. 1.7e-182;  
Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MLGVVLGALALAGLGPAPAEPPQSGQCVHEHDCFALYPGPATFLNASQICDGLRGLHLM 60

[illegible]

Search completed: December 16, 2002, 17:27:16  
Job time : 15.5 secs



entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC	EMBL; L39790; AAA74908.1; -	FT	DOMAIN	2484	2524	EGF-LIKE 42, CALCIUM-BINDING.
CC	EMBL; S69359; AAC60685.1; -	FT	DOMAIN	2525	2563	EGF-LIKE 43, CALCIUM-BINDING.
CC	HSSP; P35555; 1EMN.	FT	DOMAIN	2564	2606	EGF-LIKE 44, CALCIUM-BINDING.
DR	MGI: 95490; Fbn2.	FT	DOMAIN	2607	2646	EGF-LIKE 45, CALCIUM-BINDING.
DR	InterPro; IPR000152; ASX_hydroxyl.	FT	DOMAIN	2647	2687	EGF-LIKE 46, CALCIUM-BINDING.
DR	InterPro; IPR001881; EGF Ca.	FT	DOMAIN	2688	2727	EGF-LIKE 47, CALCIUM-BINDING.
DR	InterPro; IPR001438; EGF II.	FT	DOMAIN	115	124	BY SIMILARITY.
DR	InterPro; IPR002212; Fibril-assoc.	FT	DISULFID	119	130	BY SIMILARITY.
DR	Pfam; PF00008; EGF; 46.	FT	DISULFID	132	141	BY SIMILARITY.
DR	Pfam; PF00683; TB; 9.	FT	DISULFID	149	159	BY SIMILARITY.
DR	PRINTS; PR00010; EGFBL00D.	FT	DISULFID	153	164	BY SIMILARITY.
DR	SMART; SM00179; EGF CA; 43.	FT	DISULFID	166	175	BY SIMILARITY.
DR	SMART; SM00001; EGF-like; 3.	FT	DISULFID	180	190	BY SIMILARITY.
DR	PROSITE; PS00010; ASX_HYDROXYL; 43.	FT	DISULFID	198	207	BY SIMILARITY.
DR	PROSITE; PS00022; EGF_1; 2.	FT	DISULFID	280	292	BY SIMILARITY.
DR	PROSITE; PS01186; EGF_2; 36.	FT	DISULFID	303	316	BY SIMILARITY.
DR	PROSITE; PS01187; EGF CA; 43.	FT	DISULFID	322	334	BY SIMILARITY.
DR	Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;	FT	DISULFID	329	343	BY SIMILARITY.
KW	Repeat; Signal; Multigene family.	FT	DISULFID	345	358	BY SIMILARITY.
FT	SIGNAL	FT	DISULFID	491	503	BY SIMILARITY.
FT	CHAIN	FT	DISULFID	498	512	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	514	526	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	532	542	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	537	551	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	553	566	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	572	584	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	579	593	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	595	608	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	614	625	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	620	634	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	636	649	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	655	666	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	661	675	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	677	690	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	765	777	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	772	786	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	788	801	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	807	819	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	814	828	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	830	843	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	849	859	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	854	868	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	870	883	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	959	973	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	975	988	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1070	1082	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1077	1091	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1093	1106	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1112	1124	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1119	1133	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1135	1149	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1155	1167	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1162	1176	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1178	1191	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1197	1209	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1204	1218	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1220	1233	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1239	1250	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1246	1259	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1261	1274	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1280	1292	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1287	1301	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1303	1316	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1322	1334	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1329	1343	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1345	1358	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1364	1377	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1371	1386	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1388	1399	BY SIMILARITY.
FT	DOMAIN	FT	DISULFID	1405	1418	BY SIMILARITY.



FT DISULFID 814 827 BY SIMILARITY.  
FT DISULFID 821 836 BY SIMILARITY.  
FT DISULFID 843 856 BY SIMILARITY.  
FT DISULFID 862 875 BY SIMILARITY.  
FT DISULFID 869 884 BY SIMILARITY.  
FT DISULFID 886 899 BY SIMILARITY.  
FT DISULFID 905 917 BY SIMILARITY.  
FT DISULFID 913 926 BY SIMILARITY.  
FT DISULFID 928 941 BY SIMILARITY.  
FT DISULFID 947 956 BY SIMILARITY.  
FT DISULFID 952 965 BY SIMILARITY.  
FT DISULFID 967 980 BY SIMILARITY.  
FT DISULFID 986 998 BY SIMILARITY.  
FT DISULFID 994 1007 BY SIMILARITY.  
FT DISULFID 1009 1023 BY SIMILARITY.  
FT DISULFID 1029 1042 BY SIMILARITY.  
FT DISULFID 1036 1051 BY SIMILARITY.  
FT DISULFID 1056 1068 BY SIMILARITY.  
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 507 507 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1035 1035 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 1184 AA; 126543 MW; CA48490A55F98C5D CRC64;  
  
Query Match 12.68; Score 367.5; DB 1; Length 1184;  
Best Local Similarity 30.08; Pred. No. 4.7e-17;  
Matches 119; Conservative 29; Mismatches 124; Indels 125; Gaps 21;  
  
QY 154 EEOQCEVKAD-GFICEFHPATC-----RPLAV-----EPGAAAASVITYGTPFAARG 201  
DB 531 EQQCESNPNGPCN-HVMSLSCGEPLIVPEVRPPPEAPAPRRVS-----EAEW 582  
  
QY 202 ADFQALPVGSSNAV---APLGQLMCTAPPGAVGQHWAREAPGAWDCSVENGCGEHACNA 258  
DB 583 AGREALSLGTEAELPNLSLPGDDQDCILLPGEL-----COHLICIN 622  
  
QY 259 IPGAPRCQCPAGALQADGRSC-----TASAT-----QS 287  
DB 623 TVGSHCACFPFGLSLODDGRTRCPGHPQPEAPQEPALAKSEFSQVASTNTPLPLPQNPNT 682  
  
QY 288 CND--LCEHFCVNPNDQPGVSCMCEGYRLAADOHCEVDVDCILEPSPCP--QRCVNT 343  
DB 683 CKNGPCQKVC---STVGGSAICSCFFGYAIMADGVSCEDINECVTLHTCSRGEHCVCNT 739  
  
QY 344 QGGFECH---CYPNYDLVDGCEVEPYDPCFRANCEYQCOPLNOTSILCV-----389  
DB 740 LGSFHCYKALTCEPGYALKDGECE-EDVDEC--AMGTHTCQP-----GFLCQNTKGSFVCOA 792  
  
QY 390 ---CAEGFAPIPH-----EPHRCQMFNCNQTACPADCPNTQASCEQPEGY-I 432  
DB 793 QRRCMDGFLQDPGNCVDINECTSLSEPCRPFGFSCINTVGSYTCQNPPLI---CARGYHA 849  
  
QY 433 LDGFICTDIDECENGFCSCG---VCHNLPGTFECIC 466  
DB 850 SDDGAKCDVDNEGTGVHRCGEGQVCHNLPGSVRCDC 886  
  
RESULT 9  
FBN1\_BOVIN STANDARD; PRT; 2871 AA.  
AC FBN1\_BOVIN  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibrillin 1 precursor (MP340).  
GN FBN1.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Skin;

RA MEDLINE-95137597; PubMed-7835900;  
RA Tilstra D.J., Potter K.A., Byers P.H.;  
RT "Sequence of the coding region of the bovine fibrillin cDNA and  
RT localization to bovine chromosome 10.";  
RL Genomics 23:480-485(1994).  
RN [2]  
RP PARTIAL SEQUENCE.  
RX MEDLINE-96132851; PubMed-8557636;  
RA Gibson M.A., Hatzinikolas G., Kumaratilake J.S., Sandberg L.B.,  
RA Nicholl J.K., Sutherland G.R., Cleary E.G.;  
RT "Further characterization of proteins associated with elastic fiber  
RT microfibrils including the molecular cloning of MAGP-2 (MP25).";  
J. Biol. Chem. 271:1096-1103(1996).  
CC -|- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS  
CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE  
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.  
CC -|- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER  
CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE  
CC MICROFIBRILS.  
CC -|- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.  
CC -|- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; L28748; AAA74122.1; -;  
DR HSP; P35555; LAPJ.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF-Ca.  
DR InterPro; IPR001438; EGF-II.  
DR InterPro; IPR002212; Fibrill-assoc.  
DR Pfam; PF00008; EGF\_46.  
DR Pfam; PF00683; TB; 9.  
DR PRINTS; PR00010; EGFBL00D.  
DR SMART; SM00179; EGF\_CA; 42.  
DR SMART; SM00001; EGF\_Like; 4.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 43.  
DR PROSITE; PS00022; EGF\_1; 2.  
DR PROSITE; PS01186; EGF\_2; 38.  
DR PROSITE; PS01187; EGF\_CA; 43.  
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;  
KW Repeat; Signal; Multigene family.  
FT SIGNAL 1 27 POTENTIAL.  
FT CHAIN 28 2871 FIBRILLIN 1.  
FT DOMAIN 81 112 EGF-LIKE 1.  
FT DOMAIN 115 146 EGF-LIKE 2.  
FT DOMAIN 147 178 EGF-LIKE 3.  
FT DOMAIN 246 287 EGF-LIKE 4. CALCIUM-BINDING.  
FT DOMAIN 288 329 EGF-LIKE 5. CALCIUM-BINDING.  
FT REPEAT 330 390 TGFBP 1.  
FT DOMAIN 392 446 PRO-RICH.  
FT DOMAIN 449 489 EGF-LIKE 6.  
FT DOMAIN 490 529 EGF-LIKE 7. CALCIUM-BINDING.  
FT DOMAIN 530 571 EGF-LIKE 8. CALCIUM-BINDING.  
FT DOMAIN 572 612 EGF-LIKE 9. CALCIUM-BINDING.  
FT DOMAIN 613 653 EGF-LIKE 10. CALCIUM-BINDING.  
FT REPEAT 654 722 TGFBP 2.  
FT DOMAIN 723 764 EGF-LIKE 11. CALCIUM-BINDING.  
FT DOMAIN 765 806 EGF-LIKE 12. CALCIUM-BINDING.  
FT DOMAIN 807 846 EGF-LIKE 13. CALCIUM-BINDING.  
FT REPEAT 910 951 EGF-LIKE 14. CALCIUM-BINDING.  
FT REPEAT 952 1027 TGFBP 3.  
FT DOMAIN 1028 1069 EGF-LIKE 15. CALCIUM-BINDING.  
FT DOMAIN 1070 1112 EGF-LIKE 16. CALCIUM-BINDING.  
FT DOMAIN 1113 1154 EGF-LIKE 17. CALCIUM-BINDING.  
FT DOMAIN 1155 1196 EGF-LIKE 18. CALCIUM-BINDING.  
FT DOMAIN 1197 1237 EGF-LIKE 19. CALCIUM-BINDING.

```
FT DOMAIN 1238 EGF-LIKE 20, CALCIUM-BINDING. 1279
FT DOMAIN 1280 EGF-LIKE 21, CALCIUM-BINDING. 1321
FT DOMAIN 1322 EGF-LIKE 22, CALCIUM-BINDING. 1362
FT DOMAIN 1363 EGF-LIKE 23, CALCIUM-BINDING. 1403
FT DOMAIN 1404 EGF-LIKE 24, CALCIUM-BINDING. 1445
FT DOMAIN 1446 EGF-LIKE 25, CALCIUM-BINDING. 1486
FT DOMAIN 1487 EGF-LIKE 26, CALCIUM-BINDING. 1527
FT REPEAT 1528 TGFBP 4. 1605
FT DOMAIN 1606 EGF-LIKE 27, CALCIUM-BINDING. 1647
FT DOMAIN 1648 EGF-LIKE 28, CALCIUM-BINDING. 1688
FT REPEAT 1689 TGFBP 5. 1765
FT DOMAIN 1766 EGF-LIKE 29, CALCIUM-BINDING. 1807
FT DOMAIN 1808 EGF-LIKE 30, CALCIUM-BINDING. 1848
FT DOMAIN 1849 EGF-LIKE 31, CALCIUM-BINDING. 1890
FT DOMAIN 1891 EGF-LIKE 32, CALCIUM-BINDING. 1929
FT DOMAIN 1930 EGF-LIKE 33, CALCIUM-BINDING. 1972
FT DOMAIN 1973 EGF-LIKE 34, CALCIUM-BINDING. 2012
FT DOMAIN 2013 EGF-LIKE 35, CALCIUM-BINDING. 2054
FT REPEAT 2055 TGFBP 6. 2126
FT DOMAIN 2127 EGF-LIKE 36, CALCIUM-BINDING. 2165
FT DOMAIN 2166 EGF-LIKE 37, CALCIUM-BINDING. 2205
FT DOMAIN 2206 EGF-LIKE 38, CALCIUM-BINDING. 2246
FT DOMAIN 2247 EGF-LIKE 39, CALCIUM-BINDING. 2290
FT DOMAIN 2291 EGF-LIKE 40, CALCIUM-BINDING. 2332
FT REPEAT 2333 TGFBP 7. 2401
FT DOMAIN 2402 EGF-LIKE 41, CALCIUM-BINDING. 2443
FT DOMAIN 2444 EGF-LIKE 42, CALCIUM-BINDING. 2484
FT DOMAIN 2485 EGF-LIKE 43, CALCIUM-BINDING. 2523
FT DOMAIN 2524 EGF-LIKE 44, CALCIUM-BINDING. 2566
FT DOMAIN 2567 EGF-LIKE 45, CALCIUM-BINDING. 2606
FT DOMAIN 2607 EGF-LIKE 46, CALCIUM-BINDING. 2647
FT DOMAIN 2648 EGF-LIKE 47, CALCIUM-BINDING. 2687
FT DISULFID 85 BY SIMILARITY. 94
FT DISULFID 89 BY SIMILARITY. 100
FT DISULFID 102 BY SIMILARITY. 111
FT DISULFID 119 BY SIMILARITY. 129
FT DISULFID 123 BY SIMILARITY. 134
FT DISULFID 136 BY SIMILARITY. 145
FT DISULFID 150 BY SIMILARITY. 160
FT DISULFID 164 BY SIMILARITY. 177
FT DISULFID 168 BY SIMILARITY. 177
FT DISULFID 250 BY SIMILARITY. 262
FT DISULFID 257 BY SIMILARITY. 271
FT DISULFID 273 BY SIMILARITY. 286
FT DISULFID 292 BY SIMILARITY. 304
FT DISULFID 299 BY SIMILARITY. 313
FT DISULFID 315 BY SIMILARITY. 328
FT DISULFID 453 BY SIMILARITY. 465
FT DISULFID 460 BY SIMILARITY. 474
FT DISULFID 476 BY SIMILARITY. 488
FT DISULFID 494 BY SIMILARITY. 504
FT DISULFID 499 BY SIMILARITY. 513
FT DISULFID 515 BY SIMILARITY. 528
FT DISULFID 534 BY SIMILARITY. 546
FT DISULFID 541 BY SIMILARITY. 555
FT DISULFID 557 BY SIMILARITY. 570
FT DISULFID 576 BY SIMILARITY. 587
FT DISULFID 582 BY SIMILARITY. 596
FT DISULFID 598 BY SIMILARITY. 611
FT DISULFID 617 BY SIMILARITY. 628
FT DISULFID 623 BY SIMILARITY. 637
FT DISULFID 639 BY SIMILARITY. 652
FT DISULFID 727 BY SIMILARITY. 739
FT DISULFID 734 BY SIMILARITY. 748
FT DISULFID 750 BY SIMILARITY. 763
FT DISULFID 769 BY SIMILARITY. 781
FT DISULFID 776 BY SIMILARITY. 790
FT DISULFID 792 BY SIMILARITY. 805
FT DISULFID 811 BY SIMILARITY. 821
FT DISULFID 816 BY SIMILARITY. 830
FT DISULFID 832 BY SIMILARITY. 845
FT DISULFID 914 BY SIMILARITY. 926
FT DISULFID 921 BY SIMILARITY. 935
```

```
FT DISULFID 937 950 BY SIMILARITY.
FT DISULFID 1032 1044 BY SIMILARITY.
FT DISULFID 1039 1053 BY SIMILARITY.
FT DISULFID 1055 1068 BY SIMILARITY.
FT DISULFID 1074 1086 BY SIMILARITY.
FT DISULFID 1081 1095 BY SIMILARITY.
FT DISULFID 1097 1111 BY SIMILARITY.
FT DISULFID 1117 1129 BY SIMILARITY.
FT DISULFID 1124 1138 BY SIMILARITY.
FT DISULFID 1140 1153 BY SIMILARITY.
FT DISULFID 1159 1171 BY SIMILARITY.
FT DISULFID 1166 1180 BY SIMILARITY.
FT DISULFID 1182 1195 BY SIMILARITY.
FT DISULFID 1201 1212 BY SIMILARITY.
FT DISULFID 1208 1221 BY SIMILARITY.
FT DISULFID 1223 1236 BY SIMILARITY.
FT DISULFID 1242 1254 BY SIMILARITY.
FT DISULFID 1249 1263 BY SIMILARITY.
FT DISULFID 1265 1278 BY SIMILARITY.
FT DISULFID 1284 1296 BY SIMILARITY.
FT DISULFID 1291 1305 BY SIMILARITY.
FT DISULFID 1307 1320 BY SIMILARITY.
FT DISULFID 1326 1339 BY SIMILARITY.
FT DISULFID 1333 1348 BY SIMILARITY.
FT DISULFID 1350 1361 BY SIMILARITY.
FT DISULFID 1367 1380 BY SIMILARITY.
FT DISULFID 1374 1389 BY SIMILARITY.
FT DISULFID 1391 1402 BY SIMILARITY.
FT DISULFID 1408 1420 BY SIMILARITY.
FT DISULFID 1415 1429 BY SIMILARITY.
FT DISULFID 1431 1444 BY SIMILARITY.
FT DISULFID 1450 1461 BY SIMILARITY.
FT DISULFID 1456 1470 BY SIMILARITY.
FT DISULFID 1472 1485 BY SIMILARITY.
FT DISULFID 1491 1502 BY SIMILARITY.
FT DISULFID 1497 1511 BY SIMILARITY.
FT DISULFID 1513 1526 BY SIMILARITY.
FT DISULFID 1610 1622 BY SIMILARITY.
FT DISULFID 1617 1631 BY SIMILARITY.

Query Match 12.6%; Score 367; DB 1; Length 2871;
Best Local Similarity 31.5%; Pred. No. 1.2e-16;
Matches 111; Conservative 26; Mismatches 109; Indels 106; Gaps 20;

QY 244 DCSVENGGCEHACNAITGAPRCQCPAGAAIQADGRSCTASATSCNDLCEHFCVNP--- 300
Db 1200 ECSIMNGGCETFTCTNSEGSYECSCQPGFALMPQDRSCT-----DIDE--CEDNENIC 1249
QY 301 -----DOPGSYSCMCETGYRLAADOHRCEDVDDCILEPSPC-PQRCVNTGGFECHCYP 353
Db 1250 DGGQCTNIPGEYRCLCYDGFMASEDKMTKVDVNECDLNPNICLSGTGTCENTKGSFICHCDM 1309
QY 354 NYDLVDGE--CVPDPVDP--FRANCEYQCQPLNOT--SYLCVCAEGFA-----PIP 398
Db 1310 GYSGKKGKTGCTD-INCEIGAHNCDRHAVCTNTAGSFKSCSPGWIGDKICTDLDECS 1368
QY 399 HEPHRCQMFQNTACPADCPDNTQAS--CECPGYILDDGFICTDIDEC-EN----- 447
Db 1369 NGTHMCSQH-----ADC-KNTMGSYRCLCKEY-TGDGFTCTDLDCESENLCNGGQ 1419
QY 448 -----GGF-----CS-----GVCHNLPTGTFECICQPDLSALV 473
Db 1420 CLNAPGGYRCEDMGFVPSADGRACEDIDECSLFNICVFGTCHNLPLGFRCECEIGYELD 1479
QY 474 RHIG-----TDCDSGK-VDGGDSGSGEPPTGSGTLTPPAVGLV 513
Db 1480 RSGGNCTDVNECLDPTTCISGNCVNTPGSYTCDGCPD-----FELNPTRVGCV 1527

RESULT 10
FBNI_PIG
ID FBNI_PIG STANDARD; PRT; 2871 AA.
AC Q9TV36;
```

DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Fibrillin 1 precursor.  
GN FBN1.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Lung;  
RC MEDLINE=99156858; PubMed=10036187;  
RA Biery N.J., Eldadah Z.A., Moore C.S., Stetten G., Spencer F.,  
RA Dietz H.C.;  
RT "Revised genomic organization of FBN1 and significance for regulated  
FT gene expression.";  
RL Genomics 56:70-77(1999).  
CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS  
CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE  
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.  
CC -!- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER  
CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE  
CC MICROFIBRILS.  
CC -!- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.  
CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
-----  
DR EMBL: AF073800; AD50328.1; .  
DR HSP; P35553; IAPJ.  
DR InterPro: IPR000152; Asx\_hydroxyl.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR001881; EGF\_Ca.  
DR InterPro: IPR001438; EGF\_II.  
DR InterPro: IPR002212; Fibrill-assoc.  
DR Pfam: PF00008; EGF\_45.  
DR Pfam: PF00683; TB; 9.  
DR PRINTS; PR00010; EGFBLOOD.  
DR SMART; SM00179; EGF\_CA; 40.  
DR SMART; SN00001; EGF\_like; 6.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 41.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_2.  
DR PROSITE; PS01186; EGF\_2; 36.  
DR PROSITE; PS01187; EGF\_CA; 41.  
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;  
KW Repeat; Signal; Multigene family.  
FT SIGNAL 1 27  
FT CHAIN 28 2871 POTENTIAL.  
FT CHAIN 81 112 FIBRILLIN 1.  
FT DOMAIN 115 146 EGF-LIKE 1.  
FT DOMAIN 147 178 EGF-LIKE 2.  
FT DOMAIN 246 287 EGF-LIKE 3.  
FT DOMAIN 288 329 EGF-LIKE 4, CALCINIUM-BINDING.  
FT REPEAT 330 390 EGF-LIKE 5, CALCINIUM-BINDING.  
FT DOMAIN 392 446 TGFBP 1.  
FT DOMAIN 449 486 PRO-RICH.  
FT DOMAIN 490 529 EGF-LIKE 6.  
FT DOMAIN 530 571 EGF-LIKE 7, CALCINIUM-BINDING.  
FT DOMAIN 572 612 EGF-LIKE 8, CALCINIUM-BINDING.  
FT DOMAIN 613 653 EGF-LIKE 9, CALCINIUM-BINDING.  
FT REPEAT 654 722 EGF-LIKE 10, CALCINIUM-BINDING.  
FT DOMAIN 723 764 EGF-LIKE 11, CALCINIUM-BINDING.  
FT DOMAIN 765 806 EGF-LIKE 12, CALCINIUM-BINDING.  
FT DOMAIN 807 845 EGF-LIKE 13, CALCINIUM-BINDING.  
FT DOMAIN 910 951 EGF-LIKE 14, CALCINIUM-BINDING.  
FT REPEAT 952 1027 TGFBP 3.

```
FT DISULFID 811 821 BY SIMILARITY.
FT DISULFID 816 830 BY SIMILARITY.
FT DISULFID 832 845 BY SIMILARITY.
FT DISULFID 914 926 BY SIMILARITY.
FT DISULFID 921 935 BY SIMILARITY.
FT DISULFID 937 950 BY SIMILARITY.
FT DISULFID 1032 1044 BY SIMILARITY.
FT DISULFID 1039 1053 BY SIMILARITY.
FT DISULFID 1055 1068 BY SIMILARITY.
FT DISULFID 1074 1086 BY SIMILARITY.
FT DISULFID 1081 1095 BY SIMILARITY.
FT DISULFID 1097 1111 BY SIMILARITY.
FT DISULFID 1117 1129 BY SIMILARITY.
FT DISULFID 1124 1138 BY SIMILARITY.
FT DISULFID 1140 1153 BY SIMILARITY.
FT DISULFID 1159 1171 BY SIMILARITY.
FT DISULFID 1201 1212 BY SIMILARITY.
FT DISULFID 1208 1221 BY SIMILARITY.
FT DISULFID 1223 1236 BY SIMILARITY.
FT DISULFID 1242 1254 BY SIMILARITY.
FT DISULFID 1249 1263 BY SIMILARITY.
FT DISULFID 1265 1278 BY SIMILARITY.
FT DISULFID 1284 1296 BY SIMILARITY.
FT DISULFID 1291 1305 BY SIMILARITY.
FT DISULFID 1307 1320 BY SIMILARITY.
FT DISULFID 1326 1339 BY SIMILARITY.
FT DISULFID 1333 1348 BY SIMILARITY.
FT DISULFID 1350 1361 BY SIMILARITY.
FT DISULFID 1367 1380 BY SIMILARITY.
FT DISULFID 1374 1389 BY SIMILARITY.
FT DISULFID 1391 1402 BY SIMILARITY.
FT DISULFID 1408 1420 BY SIMILARITY.
FT DISULFID 1415 1429 BY SIMILARITY.
FT DISULFID 1450 1461 BY SIMILARITY.
FT DISULFID 1456 1470 BY SIMILARITY.
FT DISULFID 1472 1485 BY SIMILARITY.
FT DISULFID 1491 1502 BY SIMILARITY.
FT DISULFID 1497 1511 BY SIMILARITY.
FT DISULFID 1513 1526 BY SIMILARITY.
FT DISULFID 1610 1622 BY SIMILARITY.
FT DISULFID 1617 1631 BY SIMILARITY.
FT DISULFID 1633 1646 BY SIMILARITY.
FT DISULFID 1652 1663 BY SIMILARITY.
FT DISULFID 1658 1672 BY SIMILARITY.
FT DISULFID 1674 1687 BY SIMILARITY.
FT DISULFID 1770 1782 BY SIMILARITY.
FT DISULFID 1777 1791 BY SIMILARITY.
FT DISULFID 1793 1806 BY SIMILARITY.
FT DISULFID 1812 1824 BY SIMILARITY.
FT DISULFID 1818 1833 BY SIMILARITY.
FT DISULFID 1835 1847 BY SIMILARITY.
FT DISULFID 1853 1865 BY SIMILARITY.

Query Match 12.6%; Score 366; DB 1; Length 2871;
Best Local Similarity 30.6%; Pred. No. 1.3e-16;
Matches 109; Conservative 26; Mismatches 107; Indels 114; Gaps 20;

QY 244 DCSVENGCGCHACNTPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP--- 300
DB 1200 ECSIMNGGCTFTNCGSEGVSCQPGFALMPDRSCT-----DIDE--CEDNPNIC 1249
QY 301 -----DQPGSCMCGTYRLAADHRCBDDVDCTILEPSPC-PQRCVNTQGGFECHCYP 353
DB 1250 DGGQCTNIPGRCYRCLVCGFMASDKTKCTVDVNECDLNPICLSGTCENTKSGFICHCDM 1309
QY 354 NYDLVDGE--CVPVDPFC--FRANCEYQOPLNQT-SYLCVCAEFGA-----PIP 398
DB 1310 GYSGKKGKGTCTD-INCEITGAHNCORHACVNTAGSFNCSGSPGWIGDGKICTDLDECS 1368
QY 399 HEPHRCOMFCNQTACADCPNTOAS--CECPGYILDGFICTDIDEC-EN----- 447
DB 1369 NGTHMCSQH-----ADC-KNTGMYRCLCKEGY-TGDGFTCADLDECSENVKLCGNVQ 1419
```

```
QY 448 -----GGF-----CS-----GVCNHLPGTFCICGPDALV 473
DB 1420 CLYAPGGYHCEYDMGFVPSADRKSCVDSCLSNFICVFGTCHNLPGLFRCCE----- 1472
QY 474 RHIGTDCDSKVDGSGSGSEPPPS-----PTPGS-----TLTPPAVGLV 513
DB 1473 -EIGYELDRSGNCTDVNECLEPPTCISGNCVNTPGSYTCVCPDPDFELNPRVGCV 1527

RESULT 11
FBN2_HUMAN
ID FBN2_HUMAN STANDARD; PRT; 2911 AA.
AC P35556;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrillin 2 precursor.
GN FBN2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94165150; PubMed=8120105;
RA Zhang H., Apfelroth S.D., Hu W., Davis E.C., Sanguinetti C.,
RA Bonadio J., Mechem R.P., Ramirez F.;
RT "Structure and expression of fibrillin-2, a novel microfibrillar
RT component preferentially located in elastic matrices.";
RL J. Cell Biol. 124:855-863(1994).
RN [2]
RP SEQUENCE OF 752-1505 FROM N.A.
RX MEDLINE=91304567; PubMed=1852206;
RA Lee B., Godfrey M., Vitale E., Hori H., Mattel M.-G., Sarfarazi M.,
RA Tsipouras P., Ramirez F., Hollister D.;
RT "Linkage of Marfan syndrome and a phenotypically related disorder to
RT two different fibrillin genes.";
RL Nature 352:330-334(1991).
RN [3]
RP VARIANTS CCA TYR-1252 AND SER-1433, AND VARIANT ILE-964.
RX MEDLINE=96083599; PubMed=7493032;
RA Putnam E.A., Zhang H., Ramirez F., Milewicz D.M.;
RT "Fibrillin-2 (FBN2) mutations result in the Marfan-like disorder,
RT congenital contractual arachnodactyly.";
RL Nat. Genet. 11:456-458(1995).
RN [4]
RP VARIANTS CCA HIS-1114.
RX MEDLINE=98407789; PubMed=9737771;
RA Babcock D., Gasner C., Francke U., Maslen C.;
RT "A single mutation that results in an asp-to-his substitution and
RT partial exon skipping in a family with congenital contractual
RT arachnodactyly.";
RL Hum. Genet. 103:22-28(1998).
RN [5]
RP VARIANTS CCA PHE-1141 AND TRP-1252.
RX MEDLINE=20259236; PubMed=10797416;
RA Bellet S., Zhou G., Wang M., Der Kaloustian V.M., Pagon R.A.,
RA Godfrey M.;
RT "Two novel fibrillin-2 mutations in congenital contractual
RT arachnodactyly.";
RL Am. J. Med. Genet. 92:7-12(2000).
CC -!- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-2-CONTAINING MICROFIBRILS REGULATE
CC THE EARLY PROCESS OF ELASTIC FIBER ASSEMBLY.
CC -!- DISEASE: DEFECTS IN FBN2 ARE THE CAUSE OF CONGENITAL CONTRACTURAL
CC ARACHNOACTYLY (CCA) (ALSO KNOWN AS BEALS SYNDROME). CCA IS
CC PHENOTYPICALLY SIMILAR TO MARFAN SYNDROME, BUT DOES NOT EFFECT THE
CC AORTA AND THE EYES.
CC -!- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.
CC -!- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

CC	the European Bioinformatics Institute. There are no restrictions on the use of this information for any purpose, provided that the CC	
CC	use by non-profit institutions as long as its content is not modified and this statement is not removed. Usage by an	
CC	entity requires a license agreement (see <a href="http://www.isb-sib.ch">http://www.isb-sib.ch</a> ) or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> )	
CC	-----	
DR	EMBL; U03272; AAA18950.1; -	
DR	EMBL; XG2009; -; NOT_ANNOTATED_CDS.	
DR	PIR; S17063; S17063.	
DR	PIR; S31101; S31101.	
DR	HSSP; P35555; LEMN.	
DR	Genew; HGNC:3604; FBN2.	
DR	MIM; 121050; -	
DR	InterPro; IPR000152; Asx_hydroxyl.	
DR	InterPro; IPR000561; EGF-like.	
DR	InterPro; IPR001881; EGF_Ca.	
DR	InterPro; IPR001438; EGF_II.	
DR	InterPro; IPR002212; Fibrill-assoc.	
DR	Pfam; PF00008; EGF; 45.	
DR	Pfam; PF00683; TB; 9.	
DR	PRINTS; PR0010; EGFBL00D.	
DR	SMART; SM00179; EGF_CA; 43.	
DR	SMART; SM00001; EGF_Like; 3.	
DR	PROSITE; PS00010; ASX_HYDROXYL; 43.	
DR	PROSITE; PS00022; EGF_1; 2.	
DR	PROSITE; PS01186; EGF_2; 37.	
DR	PROSITE; PS01187; EGF_CA; 42.	
KW	Extracellular matrix; Calcium-binding; Glycoprotein; EGF-l	
KW	Repeat; Signal; Multigene family; Disease mutation; Polymo	
FT	SIGNAL 1 28 POTENTIAL.	
FT	CHAIN 29 2911 FIBRILLIN 2.	
FT	DOMAIN 111 142 EGF-LIKE 1.	
FT	DOMAIN 145 176 EGF-LIKE 2.	
FT	DOMAIN 176 207 EGF-LIKE 3.	
FT	DOMAIN 275 316 EGF-LIKE 4.	
FT	DOMAIN 317 358 EGF-LIKE 5, CALCIUM-BINDING.	
FT	REPEAT 359 425 TGFBP 1.	
FT	DOMAIN 493 533 EGF-LIKE 6.	
FT	DOMAIN 534 573 EGF-LIKE 7.	
FT	DOMAIN 574 615 EGF-LIKE 8.	
FT	DOMAIN 616 656 EGF-LIKE 9.	
FT	DOMAIN 657 697 EGF-LIKE 10, CALCIUM-BINDING.	
FT	REPEAT 698 766 TGFBP 2.	
FT	DOMAIN 767 808 EGF-LIKE 11.	
FT	DOMAIN 809 850 EGF-LIKE 12, CALCIUM-BINDING.	
FT	DOMAIN 851 890 EGF-LIKE 13.	
FT	DOMAIN 954 995 EGF-LIKE 14, CALCIUM-BINDING.	
FT	REPEAT 996 1071 TGFBP 3.	
FT	DOMAIN 1072 1113 EGF-LIKE 15.	
FT	DOMAIN 1114 1156 EGF-LIKE 16.	
FT	DOMAIN 1157 1198 EGF-LIKE 17.	
FT	DOMAIN 1199 1240 EGF-LIKE 18.	
FT	DOMAIN 1241 1281 EGF-LIKE 19.	
FT	DOMAIN 1282 1323 EGF-LIKE 20.	
FT	DOMAIN 1324 1365 EGF-LIKE 21.	
FT	DOMAIN 1366 1406 EGF-LIKE 22, CALCIUM-BINDING.	
FT	DOMAIN 1407 1447 EGF-LIKE 23.	
FT	DOMAIN 1448 1489 EGF-LIKE 24.	
FT	DOMAIN 1490 1530 EGF-LIKE 25.	
FT	DOMAIN 1531 1571 EGF-LIKE 26.	
FT	REPEAT 1572 1648 TGFBP 4.	
FT	DOMAIN 1649 1690 EGF-LIKE 27.	
FT	DOMAIN 1691 1732 EGF-LIKE 28.	
FT	REPEAT 1733 1806 TGFBP 5.	
FT	DOMAIN 1807 1848 EGF-LIKE 29.	
FT	DOMAIN 1849 1890 EGF-LIKE 30.	
FT	DOMAIN 1891 1932 EGF-LIKE 31.	
FT	DOMAIN 1933 1971 EGF-LIKE 32.	
FT	DOMAIN 1972 2014 EGF-LIKE 33.	
FT	DOMAIN 2015 2054 EGF-LIKE 34.	
FT	DOMAIN 2055 2096 EGF-LIKE 35.	
FT	REPEAT 2097 2169 TGFBP 6.	
FT	DOMAIN 2170 2211 EGF-LIKE 36.	
FT	DOMAIN 2212 2253 EGF-LIKE 37.	

Query Match 12.5%; Score 365; DB 1; Length 2911;  
Best Local Similarity 33.1%; Pred. No. 1.6e-16;

QY	244	DCSVENGCGCEHACNAIPGAPRCQCPAGAALQADGRSCTASATQSCNDLCHEFCVNPDP--	301
		:   :         :   :   :             :   :	
pB	1244	ECMTMNGGCTOCTNSEGYSYRSCSEGVAMPGRSCA-----DTRD--CENNPDIIC	1299





VARIANTS MFS ILE-548 AND ALA-723.  
MEDLINE-94010946; PubMed-8406497;  
Dietz H.C., McIntosh I., Sakai L.Y., Corson G.M., Chalberg S.C.,  
Peyeritz R.E., Francomano C.A.;  
"Four novel FBN1 mutations: significance for mutant transcript level  
and EGF-like domain calcium binding in the pathogenesis of Marfan  
syndrome.";  
Hum. Mol. Genet. 17:468-475(1993).  
[16]  
VARIANTS MFS SER-2144.  
MEDLINE-93278402; PubMed-8504310;  
Hewett D.R., Lynch J.R., Smith R., Sykes B.C.;  
"A novel fibrillin mutation in the Marfan syndrome which could  
disrupt calcium binding of the epidermal growth factor-like module.";  
Hum. Mol. Genet. 2:475-477(1993).  
[17]  
VARIANTS MFS R-862; V-1117; P-1137 AND F-1589, AND VARIANT A-1148.  
MEDLINE-94108431; PubMed-8281141;  
Tynan K., Comeau K., Pearson M., Wilgenbus P., Levitt D., Gasner C.,  
Berg M.A., Miller D.C., Francke U.;  
"Mutation screening of complete fibrillin-1 coding sequence: report  
of five new mutations, including two in 8-cysteine domains";  
Hum. Mol. Genet. 2:1813-1821(1993).  
[18]  
VARIANTS MFS GLY-217 AND ARG-2627.  
MEDLINE-95067970; PubMed-7977366;  
Karttunen L., Raghunath M., Loeonqvist L., Peltonen L.;  
"A compound-heterozygous Marfan patient: two defective fibrillin  
alleles result in a lethal phenotype.";  
Am. J. Hum. Genet. 55:1083-1091(1994).  
[19]  
VARIANT EL LYS-2447.  
MEDLINE-94245249; PubMed-8188302;  
Lonnqvist L., Child A., Kainulainen K., Davidson R., Puhakka L.,  
Peltonen L.;  
"A novel mutation of the fibrillin gene causing ectopia lentis.";  
Genomics 19:573-576(1994).  
[20]  
VARIANT MFS CYS-627.  
MEDLINE-94272487; PubMed-8004112;  
Hayward C., Rae A.L., Porteous M.E.M., Logie L.J., Brock L.J.;  
"Two novel mutations and a neutral polymorphism in EGF-like domains  
of the fibrillin gene (FBN1): SSCP screening of exons 15-21 in Marfan  
syndrome patients.";  
Hum. Mol. Genet. 3:373-375(1994).  
[21]  
VARIANT MFS CYS-122.  
MEDLINE-94314977; PubMed-8040326;  
Stahl-Hallengren C., Ukkonen T., Kainulainen K., Kristofersson U.,  
Saxne T., Tornqvist K., Peltonen L.;  
"An extra cysteine in one of the non-calcium-binding epidermal growth  
factor-like motifs of the FBN1 polypeptide is connected to a novel  
variant of Marfan syndrome";  
J. Clin. Invest. 94:709-713(1994).  
[22]  
VARIANT MFS TYR-1223.  
MEDLINE-94351682; PubMed-8071963;  
Hewett D.R., Lynch J.R., Child A., Sykes B.C.;  
"A new missense mutation of fibrillin in a patient with Marfan  
syndrome.";  
J. Med. Genet. 31:338-339(1994).  
[23]  
VARIANT MFS HIS-1170.  
MEDLINE-95174777; PubMed-7870075;  
Hayward C., Porteous M.E.M., Brock D.J.H.;  
"A novel mutation in the fibrillin gene (FBN1) in familial  
arachnodactyly";  
Mol. Cell. Probes 8:325-327(1994).  
[24]  
VARIANTS MFS G-217; N-1023; R-1074; Y-1242; R-1513; E-2127; W-2151;  
K-2447 AND R-2511.  
MEDLINE-94184368; PubMed-8136837;  
Kainulainen K., Karttunen L., Puhakka L., Sakai L., Peltonen L.

```

RT "Mutations in the fibrillin gene responsible for dominant ectopia
RT lentis and neonatal Marfan syndrome.";
RT Nat. Genet. 6:64-69(1994).
RN [25]
RP VARIANT SER=1127.

Query Match 12.4%; Score 363; DB 1; Length 2871;
Best Local Similarity 30.8%; Pred. No. 2.1e-16;
Matches 111; Conservative 22; Mismatches 105; Indels 122; Gaps 20;

QY 244 DCSVENGGCHACNAIGAPRCQCPAGAAQLAQRSTASATQSCNDLCHEFCVPNP--- 300
Db :||:||||| : : : : :||| : : : : :||| : : : : :|||
1200 ECSIMNGGCTFTCTNSESSEYCSQCPGALPQDRST-----DIDE--CEDPNIC 1249
QY 301 -----DQPSYSCMCETGYRLAADOHRCEDVDICILEPSPC-PQRCVNTQGGEGCHCYP 353
Db :||:||||| : : : : :||| : : : : :||| : : : : :|||
1250 DGGCTNIPGEYRCLCYDGFWASDMKTCVDVNECDLNPICLSGTCENTKGSFICHDM 1309
QY 354 NYDLVDGE--CVEPVDP--FRANCEYQCQPLNOT-SYLCVCAEGFA-----PIP 398
Db : : : : :||| : : : : :||| : : : : :||| : : : : :|||
1310 GYSCKKKTGCTD-INECEIGAHCNGKHAHVCTNAGSFKCSQSPGWIGDKIKCTDLDECS 1368
QY 399 HEPHRCOMFCNQACPADCDPNTQAS--CECPGYILDDGFICHTIDEC-EN----- 447
Db :||:||||| : : : : :||| : : : : :||| : : : : :|||
1369 NGTHMCSQH-----ADC-KNTMGYSRCLCKEY-TCDGFTCTYDLDECSENLNCNGQ 1419
QY 448 ----GGF-----CS-----GVCNHLPGTFECICGDSALV 473
Db :||:||||| : : : : :||| : : : : :||| : : : : :|||
1420 CLNAPGGYRCECDMGFVPSADGRACEDIDCSLNICVFGTCHNLPLFRCECIGYELD 1479
QY 474 RHIG-----TDCSGKVDGDSGSGEPPTPGS-----TLTPPAVGLV 513
Db :||:||||| : : : : :||| : : : : :||| : : : : :|||
1480 RSGGNCITDVNECLDPTTCISGNCVN-----TPGSYICDCCPDPFELNTRVGCV 1527

RESULT 13
FBN1_MOUSE STANDARD; PRT; 2871 AA.
ID FBN1_MOUSE 5828;
AC Q61554; Q60826;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrillin 1 precursor.
GN FBN1 OR FBN-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95130561; PubMed=7829516;
RA Yin W., Gernier J., Sanguinetti C., Smiley E., Panglilan T.,
RA Pereira L., Ramirez F., Bonadio J.;
RT "Primary structure and developmental expression of Fbn-1, the mouse
RT fibrillin gene.";
RT J. Biol. Chem. 270:1798-1806(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Kidney;
RA Ota K., Kumar A., Wada J., Liu Z., Kanwar Y.S.;
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -1- PTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER
CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE
CC MICROFIBRILS (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use.

```

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-slb.ch/announce/>  
 CC or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).  
 -----

CC	DR	EMBL; L29454; AAA56840.1; -	FT	DOMAIN	2402	2443	EGF-LIKE 41, CALCIUM-BINDING.
CC	DR	EMBL; U22493; AAA64217.1; -	FT	DOMAIN	2444	2484	EGF-LIKE 42, CALCIUM-BINDING.
CC	DR	HSP; P35555.1APJ.	FT	DOMAIN	2485	2526	EGF-LIKE 43, CALCIUM-BINDING.
CC	DR	MGD; MGI:95489; Fbn1.	FT	DOMAIN	2524	2566	EGF-LIKE 44, CALCIUM-BINDING.
CC	DR	InterPro; IPR000152; Asx_hydroxyl.	FT	DOMAIN	2567	2606	EGF-LIKE 45, CALCIUM-BINDING.
CC	DR	InterPro; IPR001881; EGF-Ca.	FT	DOMAIN	2607	2647	EGF-LIKE 46, CALCIUM-BINDING.
CC	DR	InterPro; IPR002212; Fibrin-assoc.	FT	DOMAIN	2648	2687	EGF-LIKE 47, CALCIUM-BINDING.
CC	DR	Pfam; PF00008; EGF; 46.	FT	DISULFID	85	94	BY SIMILARITY.
CC	DR	Pfam; PF00683; TB; 9.	FT	DISULFID	89	100	BY SIMILARITY.
CC	DR	SMART; SM00179; EGF-CA; 42.	FT	DISULFID	102	111	BY SIMILARITY.
CC	DR	SMART; SM00001; EGF-LIKE; 4.	FT	DISULFID	102	111	BY SIMILARITY.
CC	DR	PROSITE; PS00010; ASX_HYDROXYL; 43.	FT	DISULFID	119	129	BY SIMILARITY.
CC	DR	PROSITE; PS00022; EGF_1; 2.	FT	DISULFID	123	134	BY SIMILARITY.
CC	DR	PROSITE; PS01186; EGF_2; 38.	FT	DISULFID	136	145	BY SIMILARITY.
CC	DR	Extracellular Matrix; Calcium-binding; Glycoprotein; EGF-like domain;	FT	DISULFID	150	160	BY SIMILARITY.
CC	DR	Repeat; Signal; Multigene family.	FT	DISULFID	154	166	BY SIMILARITY.
CC	DR	POTENTIAL.	FT	DISULFID	168	177	BY SIMILARITY.
CC	DR	FIBRILLIN 1.	FT	DISULFID	250	262	BY SIMILARITY.
CC	DR	EGF-LIKE 1.	FT	DISULFID	257	271	BY SIMILARITY.
CC	DR	EGF-LIKE 2.	FT	DISULFID	273	286	BY SIMILARITY.
CC	DR	EGF-LIKE 3.	FT	DISULFID	292	304	BY SIMILARITY.
CC	DR	EGF-LIKE 4.	FT	DISULFID	299	313	BY SIMILARITY.
CC	DR	EGF-LIKE 5.	FT	DISULFID	315	328	BY SIMILARITY.
CC	DR	EGF-LIKE 6.	FT	DISULFID	453	465	BY SIMILARITY.
CC	DR	EGF-LIKE 7.	FT	DISULFID	460	474	BY SIMILARITY.
CC	DR	EGF-LIKE 8.	FT	DISULFID	476	488	BY SIMILARITY.
CC	DR	EGF-LIKE 9.	FT	DISULFID	494	504	BY SIMILARITY.
CC	DR	EGF-LIKE 10.	FT	DISULFID	499	513	BY SIMILARITY.
CC	DR	EGF-LIKE 11.	FT	DISULFID	515	528	BY SIMILARITY.
CC	DR	EGF-LIKE 12.	FT	DISULFID	534	546	BY SIMILARITY.
CC	DR	EGF-LIKE 13.	FT	DISULFID	541	555	BY SIMILARITY.
CC	DR	EGF-LIKE 14.	FT	DISULFID	557	570	BY SIMILARITY.
CC	DR	EGF-LIKE 15.	FT	DISULFID	576	587	BY SIMILARITY.
CC	DR	EGF-LIKE 16.	FT	DISULFID	582	596	BY SIMILARITY.
CC	DR	EGF-LIKE 17.	FT	DISULFID	598	611	BY SIMILARITY.
CC	DR	EGF-LIKE 18.	FT	DISULFID	617	628	BY SIMILARITY.
CC	DR	EGF-LIKE 19.	FT	DISULFID	623	637	BY SIMILARITY.
CC	DR	EGF-LIKE 20.	FT	DISULFID	639	652	BY SIMILARITY.
CC	DR	EGF-LIKE 21.	FT	DISULFID	727	739	BY SIMILARITY.
CC	DR	EGF-LIKE 22.	FT	DISULFID	734	748	BY SIMILARITY.
CC	DR	EGF-LIKE 23.	FT	DISULFID	750	763	BY SIMILARITY.
CC	DR	EGF-LIKE 24.	FT	DISULFID	769	781	BY SIMILARITY.
CC	DR	EGF-LIKE 25.	FT	DISULFID	776	790	BY SIMILARITY.
CC	DR	EGF-LIKE 26.	FT	DISULFID	792	805	BY SIMILARITY.
CC	DR	EGF-LIKE 27.	FT	DISULFID	811	821	BY SIMILARITY.
CC	DR	EGF-LIKE 28.	FT	DISULFID	816	830	BY SIMILARITY.
CC	DR	EGF-LIKE 29.	FT	DISULFID	832	845	BY SIMILARITY.
CC	DR	EGF-LIKE 30.	FT	DISULFID	914	926	BY SIMILARITY.
CC	DR	EGF-LIKE 31.	FT	DISULFID	921	935	BY SIMILARITY.
CC	DR	EGF-LIKE 32.	FT	DISULFID	937	950	BY SIMILARITY.
CC	DR	EGF-LIKE 33.	FT	DISULFID	1032	1044	BY SIMILARITY.
CC	DR	EGF-LIKE 34.	FT	DISULFID	1039	1053	BY SIMILARITY.
CC	DR	EGF-LIKE 35.	FT	DISULFID	1055	1068	BY SIMILARITY.
CC	DR	EGF-LIKE 36.	FT	DISULFID	1074	1086	BY SIMILARITY.
CC	DR	EGF-LIKE 37.	FT	DISULFID	1081	1095	BY SIMILARITY.
CC	DR	EGF-LIKE 38.	FT	DISULFID	1097	1111	BY SIMILARITY.
CC	DR	EGF-LIKE 39.	FT	DISULFID	1117	1129	BY SIMILARITY.
CC	DR	EGF-LIKE 40.	FT	DISULFID	1124	1138	BY SIMILARITY.
CC	DR	EGF-LIKE 41.	FT	DISULFID	1140	1153	BY SIMILARITY.
CC	DR	EGF-LIKE 42.	FT	DISULFID	1159	1171	BY SIMILARITY.
CC	DR	EGF-LIKE 43.	FT	DISULFID	1166	1180	BY SIMILARITY.
CC	DR	EGF-LIKE 44.	FT	DISULFID	1182	1195	BY SIMILARITY.
CC	DR	EGF-LIKE 45.	FT	DISULFID	1201	1212	BY SIMILARITY.
CC	DR	EGF-LIKE 46.	FT	DISULFID	1208	1221	BY SIMILARITY.
CC	DR	EGF-LIKE 47.	FT	DISULFID	1223	1236	BY SIMILARITY.
CC	DR	EGF-LIKE 48.	FT	DISULFID	1242	1254	BY SIMILARITY.
CC	DR	EGF-LIKE 49.	FT	DISULFID	1249	1263	BY SIMILARITY.
CC	DR	EGF-LIKE 50.	FT	DISULFID	1265	1278	BY SIMILARITY.
CC	DR	EGF-LIKE 51.	FT	DISULFID	1284	1296	BY SIMILARITY.
CC	DR	EGF-LIKE 52.	FT	DISULFID	1291	1305	BY SIMILARITY.
CC	DR	EGF-LIKE 53.	FT	DISULFID	1307	1320	BY SIMILARITY.
CC	DR	EGF-LIKE 54.	FT	DISULFID	1326	1339	BY SIMILARITY.
CC	DR	EGF-LIKE 55.	FT	DISULFID	1333	1348	BY SIMILARITY.
CC	DR	EGF-LIKE 56.	FT	DISULFID	1350	1361	BY SIMILARITY.



```

FT DISULFID 372 385 BY SIMILARITY.
FT DISULFID 387 400 BY SIMILARITY.
FT DISULFID 406 417 BY SIMILARITY.
FT DISULFID 413 426 BY SIMILARITY.
FT DISULFID 428 441 BY SIMILARITY.
FT DISULFID 447 458 BY SIMILARITY.
FT DISULFID 454 467 BY SIMILARITY.
FT DISULFID 469 482 BY SIMILARITY.
FT DISULFID 488 499 BY SIMILARITY.
FT DISULFID 495 508 BY SIMILARITY.
FT DISULFID 510 523 BY SIMILARITY.
FT DISULFID 529 540 BY SIMILARITY.
FT DISULFID 536 549 BY SIMILARITY.
FT DISULFID 551 564 BY SIMILARITY.
FT DISULFID 570 581 BY SIMILARITY.
FT DISULFID 577 590 BY SIMILARITY.
FT DISULFID 592 605 BY SIMILARITY.
FT DISULFID 611 622 BY SIMILARITY.
FT DISULFID 618 631 BY SIMILARITY.
FT DISULFID 633 646 BY SIMILARITY.
FT CARBOHYD 221 221 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 890 890 N-LINKED (GLCNAC. .) (POTENTIAL).
FT VARSPLIC 861 879 MISSING (IN SHORT ISOFORM).
FT CONFLICT 356 356 K -> E (IN REF. 3).
FT CONFLICT 394 394 V -> E (IN REF. 3).
FT CONFLICT 644 644 R -> E (IN REF. 2).
FT CONFLICT 755 755 F -> L (IN REF. 3).
FT CONFLICT 935 935 L -> F (IN REF. 2).
SQ SEQUENCE 956 AA; 106840 MW; 826B7F347178FC80 CRC64;

Query Match 12.3%; Score 358; DB 1; Length 956;
Best Local Similarity 31.6%; Pred. No. 1.6e-16;
Matches 86; Conservative 41; Mismatches 101; Indels 44; Gaps 15;

QY 245 CSVENGGCEHACNAIFGAPRCQCPAGAAIQAQGRSCTA-SATQSCMDLCEHFCVNPDPQ 303
Db 283 CAMEDHNCQLQVNVPGSFVQCYSYALAEKRCVAVDYCASNHGECV NAD-- 339
QY 304 GSYSCMCTGYRLAADQHRCEVDVDCILRPSPCQRCVNTQGGFECYVNYDL-VDGR 362
Db 340 GSYLCQCHGFALNPDKTKTKIDYCASSNHGQHCQHCVNTDSDSYCHCLKGLFTLPDKKT 399
QY 363 VEPVDPGF--RANCETQCOPLNQTSLYLCVCAEGFAPIPH-----EPHRCQMEC 408
Db 400 CRRINYCALNKPCEHCVNMEE-SYICRCHRGYTLDPNGKTCRSVDHCAQQDHGCEQLC 458
QY 409 NQACPADCDPNTQAS--CECPGYTLDDGF-ICTDIDFC---ENGFGCSGVCHNLPGRF 462
Db 459 -----LNTDFSVQCSEGLFNLDELKTCRSVDYCLLSDRG--CEYSCVNDRSF 506
QY 463 ECICGPDLSALVRHIGTDCSGKVDG---GDSG 491
Db 507 ACQC-PEGHVLRSDGKTC--AKLDSALGDBG 535

RESULT 15
FBL2_MOUSE
ID FBL2_MOUSE STANDARD; PRT; 1221 AA.
AC P37889; Q9MU12;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibulin-2 precursor.
GN FBLN2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-35.
RC TISSUE=Fibroblast;
RX MEDLINE=94064787; PubMed=8245130;
RA Pan T.-C., Sasaki T., Zhang R.-Z., Faessler R., Timpl R., Chu M.-L.;

```

```

RT "Structure and expression of fibulin-2, a novel extracellular matrix
RT protein with multiple EGF-like repeats and consensus motifs for
RT calcium binding.";
RL J. Cell Biol. 123:1269-1277(1993).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=9337686; PubMed=10406956;
RA Graessel S., Sicot F.-A., Gotta S., Chu M.-L.;
RT "Mouse fibulin-2 gene. Complete exon-intron organization and promoter
RT characterization.";
RL Eur. J. Biochem. 263:471-477(1999).
CC -1- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS
CC CALCIUM DEPENDENT.
CC -1- SUBUNIT: HOMOTRIMER; DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: Extracellular matrix.
CC -1- ALTERNATIVE PRODUCTS: At least 2 isoforms; 1 (shown here) and
CC 2/EGF3-less; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: COMPONENT OF BOTH BASEMENT MEMBRANES AND OTHER
CC CONNECTIVE TISSUES.
CC -1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X75285; CAA53040.1; -.
DR EMBL; AF135253; AAD34456.1; JOINED.
DR EMBL; AF135239; AAD34456.1; JOINED.
DR EMBL; AF135240; AAD34456.1; JOINED.
DR EMBL; AF135241; AAD34456.1; JOINED.
DR EMBL; AF135242; AAD34456.1; JOINED.
DR EMBL; AF135243; AAD34456.1; JOINED.
DR EMBL; AF135244; AAD34456.1; JOINED.
DR EMBL; AF135245; AAD34456.1; JOINED.
DR EMBL; AF135246; AAD34456.1; JOINED.
DR EMBL; AF135247; AAD34456.1; JOINED.
DR EMBL; AF135248; AAD34456.1; JOINED.
DR EMBL; AF135249; AAD34456.1; JOINED.
DR EMBL; AF135250; AAD34456.1; JOINED.
DR EMBL; AF135251; AAD34456.1; JOINED.
DR EMBL; AF135252; AAD34456.1; JOINED.
DR PIR; A49457; A49457.
DR HSP; P00736; 1APO.
DR MGD; MGI:95488; Fbln2.
DR InterPro; IPR000020; Anaphylatoxin.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR01881; EGF_Ca.
DR Pfam; PF00008; EGF; 6.
DR Pfam; PF01821; ANATO; 2.
DR SMART; SM00104; ANATO; 3.
DR SMART; SM00179; EGF_CA; 9.
DR SMART; SM00001; EGF_like; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; 5.
DR PROSITE; PS01177; ANAPHYLATOXIN_1; 3.
DR PROSITE; PS01178; ANAPHYLATOXIN_2; 3.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 5.
DR PROSITE; PS01187; EGF_CA; 10.
KW Signal; Glycoprotein; Extracellular matrix; Plasma; EGF-like domain;
FT SIGNAL 1 26
FT CHAIN 27 1221 FIBULIN-2.
FT DOMAIN 27 434 N.
FT DOMAIN 27 176 SUBDOMAIN NA (CYS-RICH).
FT DOMAIN 177 434 SUBDOMAIN NB (CYS-FREE).
FT DOMAIN 435 477 ANAPHYLATOXIN-LIKE 1.
FT DOMAIN 478 510 ANAPHYLATOXIN-LIKE 2.

```

DOMAIN	511	543	ANAPHYLATOXIN-LIKE 3.
FT	594	635	EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
FT	669	708	EGF-LIKE 2.
FT	709	755	EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT	756	800	EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT	801	846	EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT	847	894	EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT	895	937	EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT	938	979	EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT	980	1018	EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT	1019	1061	EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).
FT	1062	1106	EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
FT	1111	1221	DOMAIN III.
FT	421	423	CELL ATTACHMENT SITE (POTENTIAL).
FT	435	462	BY SIMILARITY.
FT	436	469	BY SIMILARITY.
FT	449	470	BY SIMILARITY.
FT	479	508	BY SIMILARITY.
FT	492	509	BY SIMILARITY.
FT	511	535	BY SIMILARITY.
FT	512	542	BY SIMILARITY.
FT	525	543	BY SIMILARITY.
FT	598	610	BY SIMILARITY.
FT	606	619	BY SIMILARITY.
FT	621	634	BY SIMILARITY.
FT	673	683	BY SIMILARITY.
FT	679	692	BY SIMILARITY.
FT	694	707	BY SIMILARITY.
FT	713	726	BY SIMILARITY.
FT	720	735	BY SIMILARITY.
FT	742	754	BY SIMILARITY.
FT	805	818	BY SIMILARITY.
FT	812	827	BY SIMILARITY.
FT	833	845	BY SIMILARITY.
FT	899	912	BY SIMILARITY.
FT	906	921	BY SIMILARITY.
FT	923	936	BY SIMILARITY.
FT	942	954	BY SIMILARITY.
FT	950	963	BY SIMILARITY.
FT	965	978	BY SIMILARITY.
FT	984	993	BY SIMILARITY.
FT	1009	1002	BY SIMILARITY.
FT	1004	1017	BY SIMILARITY.
FT	1023	1035	BY SIMILARITY.
FT	1031	1044	BY SIMILARITY.
FT	1046	1060	BY SIMILARITY.
FT	1066	1079	BY SIMILARITY.
FT	1073	1088	BY SIMILARITY.
FT	1093	1105	BY SIMILARITY.
FT	179	179	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	497	497	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	737	737	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	1072	1072	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	709	755	MISSING (IN ISOFORM 2).
FT	140	159	HSGRKYAAGHIVHLSRAC -> TVAVSICWVYRPLILP
FT	348	348	GF (IN REF. 2).
FT	507	507	S -> L (IN REF. 2).
FT	1102	1102	Q -> QQ (IN REF. 2).
FT	1221	131818	MM; 87DB2A10A9FDc45F CRC64
FT	1221	131818	AA; 87DB2A10A9FDc45F CRC64
FT	348	348	S -> L (IN REF. 2).
FT	507	507	Q -> QQ (IN REF. 2).
FT	1102	1102	Q -> E (IN REF. 2).
FT	1221	131818	MM; 87DB2A10A9FDc45F CRC64
FT	1221	131818	AA; 87DB2A10A9FDc45F CRC64
FT	348	348	S -> L (IN REF. 2).
FT	507	507	Q -> QQ (IN REF. 2).
FT	1102	1102	Q -> E (IN REF. 2).
FT	1221	131818	MM; 87DB2A10A9FDc45F CRC64
FT	1221	131818	AA; 87DB2A10A9FDc45F CRC64
FT	348	348	S -> L (IN REF. 2).
FT	507	507	Q -> QQ (IN REF. 2).
FT	1102	1102	Q -> E (IN REF. 2).
FT	1221	131818	MM; 87DB2A10A9FDc45F CRC64
FT	1221	131818	AA; 87DB2A10A9FDc45F CRC64
FT	348	348	S -> L (IN REF. 2).
FT	507	507	Q -> QQ (IN REF. 2).
FT	1102	1102	Q -> E (IN REF. 2).
FT	1221	131818	MM; 87DB2A10A9FDc45F CRC64
FT	1221	131818	AA; 87DB2A10A9FDc45F CRC64
FT	348	348	S -> L (IN REF. 2).
FT	507	507	Q -> QQ (IN REF. 2).
FT	1102	1102	Q -> E (IN REF. 2).
FT	1221	131818	MM; 87DB2A10A9FDc45F CRC64
FT	1221	131818	AA; 87DB2A10A9FDc45F CRC64
FT	348	348	S -> L (IN REF. 2).
FT	507	507	Q -> QQ (IN REF. 2).
FT	1102	1102	Q -> E (IN REF. 2).
FT	1221	131818	MM; 87DB2A10A9FDc45F CRC64
FT	1221	131818	AA; 87DB2A10A9FDc45F CRC64
FT	348	348	S -> L (IN REF. 2).
FT	507	507	Q -> QQ (IN REF. 2).
FT	1102	1102	Q -> E (IN REF. 2).

Qy	230	AVQCHWARE--APGNWDCSV-----ENGCEHACNAIPQAPRCQCPAGAA	272
Db	640	APQLDTAREAPRESAQVSPNTIPLVPVPQNTCKNGPCRCQRVVGDATMCSCFPFGYA	699
Qy	273	LQADGRSC-----TASATQSCNDLCEHFCVPNPDPQSGYSYC-----MCETGYRLAADQR	322
Db	700	IMADGVSCDQDECLMGTDHCS--WKQFCV---NTLGSFYCVNHTVLCAEGYILNA--HRK	753
Qy	323	CEVDVDCILLESPC--PQRCVNTQGGFECH-----CYPNTYDLVBGCEVPVDPFCFRA--NC	374
Db	754	CVDINECVTLDTCTRAEHCVNTPGSGQCYKALCTCFGYVLTGDGCTD-VDECVTGTGHC	812
Qy	375	E--YOCOPLNTQSYLCV-----CAEGFAPIPH-----EPRHQCMFCNQTACPA	415
Db	813	QAGFSQCN-YKGSFYQARQCHMGDFLQDPNGNCVDINECTSLLEPCRSGFSCINTVGSY	871
Qy	416	DCDPNTQASCECPGY-ILDDGFTCTDIDECENGFCSG---YCHNLPGTFTECIGP	468
Db	872	TCORNPWL---CGRGYHANBEGSCVDVNECTGYVHRCGEQLCYLNPUSYRCDCKP	925

Search completed: December 16, 2002, 17:25:00  
Job time : 20.5 secs





the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
```

---

```

CC CC EMBL; AF074856; AAC63274.1; -
CC CC EMBL; AF081789; AAC62649.1; -
CC CC EMBL; AF099939; AAD47906.1; -
CC CC EMBL; AF099938; AAD47906.1; JOINED.
CC CC HSSP; P35555; IEMN.
CC CC MGD; MGI:106664; Ly68.
CC CC InterPro; IPR000152; Asx_hydroxyl.
CC CC InterPro; IPR000561; EGF-like.
CC CC InterPro; IPR001881; EGF-Ca.
CC CC InterPro; IPR001304; Lectin_C.
CC CC Pfam; PF00008; EGF; 5.
CC CC Pfam; PF00059; Lectin_c; 1.
CC CC SMART; SM00034; CLECT; 1.
CC CC SMART; SM00179; EGF_Ca; 3.
CC CC SMART; SM00001; EGF_like; 2.
CC CC PROSITE; PS00010; ASX_HYDROXYL; 3.
CC CC PROSITE; PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
CC CC PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
CC CC PROSITE; PS01186; EGF_2; 3.
CC CC PROSITE; PS01187; EGF_Ca; 3.
CC CC Receptor; EGF-like domain; Signal; Transmembrane; Glycoprotein;
CC CC Repeat; Lectin.
CC CC SIGNAL 1 22 POTENTIAL.
CC CC CHAIN 23 644 COMPLEMENT COMPONENT C1Q RECEPTOR.
CC CC DOMAIN 23 572 EXTRACELLULAR (POTENTIAL).
CC CC TRANSMEM 573 593 POTENTIAL.
CC CC DOMAIN 594 644 CYTOPLASMIC (POTENTIAL).
CC CC DOMAIN 31 173 C-TYPE LECTIN.
CC CC DOMAIN 257 298 EGF-LIKE 1.
CC CC DOMAIN 299 341 EGF-LIKE 2.
CC CC DOMAIN 342 381 EGF-LIKE 3.
CC CC DOMAIN 382 423 EGF-LIKE 4.
CC CC DOMAIN 424 465 EGF-LIKE 5.
CC CC DISULFID 261 272 BY SIMILARITY.
CC CC DISULFID 268 282 BY SIMILARITY.
CC CC DISULFID 284 297 BY SIMILARITY.
CC CC DISULFID 303 314 BY SIMILARITY.
CC CC DISULFID 308 325 BY SIMILARITY.
CC CC DISULFID 327 340 BY SIMILARITY.
CC CC DISULFID 346 355 BY SIMILARITY.
CC CC DISULFID 351 364 BY SIMILARITY.
CC CC DISULFID 366 380 BY SIMILARITY.
CC CC DISULFID 386 397 BY SIMILARITY.
CC CC DISULFID 393 406 BY SIMILARITY.
CC CC DISULFID 408 422 BY SIMILARITY.
CC CC DISULFID 428 440 BY SIMILARITY.
CC CC DISULFID 436 449 BY SIMILARITY.
CC CC DISULFID 451 464 BY SIMILARITY.
CC CC CARBOHYD 102 102 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CC CARBOHYD 322 322 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CC SEQUENCE 644 AA; 69354 MW; EB4351648BF8635A CRC64;

Query Match 18.6%; Score 543; DB 1; Length 644;
Best Local Similarity 30.0%; Pred. No. 7.9e-29;
Matches 167; Conservative 55; Mismatches 211; Indels 124; Gaps 27;

QY 5 LVILGALAGLPAPAEQPGGSCQVHDFCALYPGPATFLNASTQICDGLRGLMTVRS 64
DB 9 LLILGLGQPWAG--AAADSQ--AVVCEGTACTAHHWGLKLSABAQAHRNNGNGLATVRS 64

QY 65 SVAA-----DVISLLNGDGVGRR--RLMTIGLQLPPCGDPRKRLGFRQWTVGNNTS 118
DB 65 EEARHVQQAQLTLLTKAPLEAKMKGFWIGLORENGNCTYHDL-PMRGFSWVGGEDTA 123

QY 119 YSKWARLDLNGAPLCGLPVASAAATVPSE-PIWEEQOCEV-----KADGFLCEFFH 171
DB 124 YSNWYKASKSSCFFKRCVSLILDLSLTPHPLPKWHESPCTPEAPGNSIEGFLCKNF 183
```

---

```

QY 172 PATCRPLAV-EPGAAAAVSIITYGTFPAARGADFOALPVGSSAAVAPLGLQ-----LM 223
DB 184 KGMCRPLALGGPG-----RVITYTTPFOATTSSLEAVPPASVANVA-CGDEAKSETHYFL 236
QY 224 CT-APPGAVQGHAREAP-----GAWDCSVENGSCHEHAC-NAIPGAPRCPCPAGAAQAQDG 277
DB 237 CNEKTPGIF--HWSSGGLCVSPKFGCSFNNGGQQDCFEFGDGSFRCGRFRLDDL 294
QY 278 RSCATASQSCNDLCEHFCVNPDPG-----SYSCMCETGYRLAADQHRCEVD 327
DB 295 VTCAS-----RNPSSNPCTGGGMCHSVPLSENTCRPCSYQLDSSVHCVCDID 344
QY 328 DCILEPSPCQRCVNTQGGFECHCYPNYDLVG---ECVEPVDPCFRANCEYQCQPLNQ 384
DB 345 EC-QDSPCAQDCVNTLGSFHCCEWVGYO-PSGPKEACEDVDECAANSP----- 392
QY 385 SYLCVCAEAGFAPIPHEPHRCQMFQNTACPADCDPNTQAS--CECEGYIL--DDGFICT 440
DB 393 -----CAQGC-----NTDGSFYCSCKEGYIVSGEDSTQCE 423
QY 441 DIDECEN--GGFCSGVCHNLPGTFECICGPDLSALVRHIGTDCDSG-----KVD 486
DB 424 DIDECSDARGNFCDSLFCNTDGSFRCGCPGWELAPN-GVFCSRGTVFSELPARPPQKED 482
QY 487 GGDGSGSEPPSPPTGS 503
DB 483 NDRKRESTMPTTEMPSS 499

RESULT 6
CD93_RAT STANDARD; PRT; 643 AA.
AC Q9ET61; Q9JIZ6;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Complement component C1q receptor precursor (Complement component 1, q
DE subcomponent, receptor 1) (ClqR) (ClqR(p)) (Clq/MBL/SPA receptor)
DE (CD93 antigen) (Cell surface antigen AA4).
GN ClQ1 OR CD93 OR ClQRP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PVG; TISSUE=Natural killer cells;
RX MEDLINE-20545218; PubMed-11093152;
RA Lovik G., Vaage J.F., Dissen E., Szpirer C., Ryan J.C., Rolstad B.;
RT "Characterization and molecular cloning of rat ClqR, a receptor on NK
RT cells.";
RL Eur. J. Immunol. 30:3355-3362(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Lung;
RX MEDLINE-20507883; PubMed-10934210;
RA Dean Y.D., McGreal E.P., Akatsu H., Gasque P.;
RT "Molecular and cellular properties of the rat Aa antigen, a C-type
RT lectin-like receptor with structural homology to thrombomodulin.";
RL J. Biol. Chem. 275:34382-34392(2000).
CC -I- FUNCTION: Receptor (or element of a larger receptor complex) for
CC C1q, mannose-binding lectin (MBL2) and pulmonary surfactant
CC protein A (SPA). May mediate the enhancement of phagocytosis in
CC monocytes and macrophages upon interaction with soluble defense
CC collagens. May play a role in intercellular adhesion.
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- TISSUE SPECIFICITY: Widely expressed. Highly expressed in lung and
CC heart. Expressed at lower level in brain, thymus, liver, spleen,
CC intestine, kidney, adrenal gland, muscle and testis. Expressed on
CC endothelial cells, platelets, undifferentiated monocytes and
CC circulating natural killer cells.
CC -I- PTM: N- and O-glycosylated (By similarity).
```



[illegible]

```

Query Match      20.0%; Score 583; DB 1; Length 652;
Best Local Similarity 31.8%; Pred. No. 1.9e-31;
Matches 181; Conservative 55; Mismatches 210; Indels 124; Gaps 32;

QY      2  LGVILVGLALALA--GLGFPAPEPQPGSGOCVEHDCFALYPGPATFNLASQICDGLRGHL 59
      :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      5  MGLLLLLLLLTOPGAGTGADTE---AVVCVGTCACYTAHSGKLSAAEAQHNCNQNGNL 60

QY     60  MTRVSSVAAD---VISLLNGDGVGRR--RLWIGLQLPPG--CGDKPRKGLPLRGFQWVT 112
      :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     61  ATVKSKEEAQHVRVLAQLLRRAAULTARMSKEWIGLQREKGKCLDPSL--PLKGFSSWG 118
      :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY     113  GDNNITSRWARLDLNCAPLCGPCLVA--VSAEAATVPSE-PIWEEQOC-----EYKAD 163
      |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     119  GGEDPTYSNNHKELRNSC--ISKRCVSLLLDLQPLPLSRLPKWSGEGPCGSPGSPGSNIE 176

QY     164  GFLCEHFHAPCRPLAV-EPGAAAAVSITYGTPFAARGADFOALPGVSSAANAAPLGLQL 222
      :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     177  GFVCKSFSGKNCRLALGGPG-----QVTTTPTFTQTTSSLEAVPFAANA'V----- 224

QY     223  MC-TAPPGAVQGHW--AREAPGAWD-----CSVENGGCEHAC-NAIPCAP 263
      :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     225  -CGEGKDEFTQSHYFLCKERAPVFDWSSGPLCVSPKYGCNFNNGGCHQDCFEGBGGSF 283

QY     264  RCOCPGAALQADGRSATSATOSCNLDCE--HFCVNPDPQSGYSXCMCTEYFLAADQH 321
      :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     284  LCGCRPGFRLDLLVTC-ASRNCSPSSPCRGGATCVLGP-HGKNYTCRCQPGYQLDSSQL 341

QY     322  RCEVDVDCILEPSPCQRQCVNTGGGECHCYPNYDLVDGECVEPDPFCFRANCEYQCOPL 381
      :|:|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     342  DCVDVDEEC--QDSPCAQECVNTPGGFRCECVWG'-----EPGGP-----GEGACQDV 386

QY     382  NQTSYLCVCAEGEAPIPHEPHRCQMFNCNQTAACADCPDNTQAS--CBCPBGYIL--DDGF 437
      :|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     387  DE-----CALGRSP-----CAQGC--TNRDGSFHCSEGYVLAGEDGT 423

QY     438  ICTDIDEC--ENGFGSGVCHNLPGTPECICGPDLSALVRHIGTDCDSGV-----D 486
      :|:|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db     424  QCQDVDECVGPGCLDLSLCTQSGFHCGLCGPWLA'PN-GVSC'CTMGVPSYSLGPPSPGPD 482

```

QY 487 GGDGSGSE-----PPSPTPGSTTTPPA 509  
| | | | | | | | | |  
Db 483 EEDKGEREGSTVPRAATASPTRGECTPKA 512

RESULT 5

ID	CD93_MOUSE	STANDARD;	PRT;	644 AA.
AC	089103;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DE	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Complement component C1q receptor precursor (Complement component 1, q			
DE	subcomponent, receptor 1) (ClqrP) (C1qR(p)) (C1q/MBL/SPA receptor)			
DE	(CD93 antigen), (Cell surface antigen A44) (lymphocyte antigen 68).			
GN	C1QR1 OR CD93 OR C1QRP OR Lf68 OR A44.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBL_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129/Sv;			
RX	PubMed=1074255;			
RA	Kim T.S., Park M., Nepomuceno R.R., Palmirini G., Winokur S.,			
RA	Cotman C.A., Bengtsson U., Tenner A.J.;			
RT	"Characterization of the murine homolog of ClQR(P): identical cellular			
RT	expression pattern, chromosomal location and functional activity of			
RT	the human and murine ClQR(P).";			
RL	Mol. Immunol. 37:377-389(2000).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Leukemia;			
RX	PubMed=10403644;			
RA	Petrenko O., Beavis A., Klaine M., Kittappa R., Godin I.,			
RA	Lemischka I.R.;			
RT	"The molecular characterization of the fetal stem cell marker AA4.";			
RL	Immunity 10:691-700(1999).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129/Sv; TISSU=Spleen, and Endothelial cells;			
RX	MEDLINE=99359842; PubMed=10430665;			
RA	Norsworthy P.J., Taylor P.R., Walport M.J., Botto M.;			
RT	"Cloning of the mouse homolog of the 126-kDa human C1q/MBL/SP-A			
RT	receptor, ClqrP.";			
RL	Mamm. Genome 10:789-793(1999).			
CC	-I- FUNCTION: Receptor (or element of a larger receptor complex) for			
CC	C1q, mannose-binding lectin (MBL2) and pulmonary surfactant			
CC	protein A (SPA). May mediate the enhancement of phagocytosis in			
CC	monocytes and macrophages upon interaction with soluble defense			
CC	collagens. May play a role in intercellular adhesion. Marker for			
CC	early multipotent hematopoietic precursor cells. May play a role			
CC	in cell-cell interactions during hematopoietic and vascular			
CC	development.			
CC	-I- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-I- TISSUE SPECIFICITY: Expressed in lung, heart and bone marrow.			
CC	Expressed at lower level in ovary, whole embryo and fetal liver.			
CC	Not detected in brain, adult liver or thymus. Highly expressed in			
CC	peritoneal cavity and bone marrow macrophages. Not detected in			
CC	epithelial cells.			
CC	-I- DEVELOPMENTAL STAGE: First detectable in day 9 embryos, in the			
CC	endocardium and vascular endothelium in the anterior part of the			
CC	embryo. Expression in endothelial cells, initially restricted to			
CC	aorta, omphalomesenteric and umbilical arteries, later extends to			
CC	subcardinal veins, intersomatic arteries and perimeural vessels.			
CC	On day 10, detectable in the entire embryo.			
CC	-I- PTM: N- and O-glycosylated (By similarity).			
CC	-I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.			
CC	-I- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation

```

Db 123 CHCDTGELVDEGVDPVDFNCEYQCPVGRSEHKICAGFAVPVPGAPHKQWFC 182
Qy 409 NOTACPADCPNTQASCEPGYILDDGFICTDIDECENGFCSCVCHNLPTGTEFCICGP 468
Db 183 NOTSCPADCPHYTICRPGYIIDEGSTCTDINECDT-NICPGQCHNLPTGTEFCICGP 241
Qy 469 DSALVRHGTDCDGSKV-----DGGDSGSGRPPPTGSLTLP-PA-VGLVHSG 516
Db 242 DSALSGQIGIDCDPTQVNEERGTPEDYG--GSGEPVPTPGATARPSPAPAGPLHSG 297

RESULT 4
ID CD93_HUMAN STANDARD; PRT; 652 AA.
AC Q9NPY3; O00274;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Complement component C1q receptor precursor (Complement component 1, q
DE subcomponent, receptor 1) (C1qR) (C1qR(p)) (C1q/MBL/SPA receptor)
DE (CD93 antigen) (CDw93).
GN C1QR1 OR CD93.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX PubMed=9047234;
RA Nepomuceno R.R., Henschen-Edman A.H., Burgess W.H., Tenner A.J.;
RT "cDNA cloning and primary structure analysis of C1qR(p), the human
RT C1q/MBL/SPA receptor that mediates enhanced phagocytosis in vitro.";
RL Immunity 6:119-129(1997).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT ALA-318.
RX PubMed=11781389;
RA Steinberger P., Szekeres A., Willie S., Stockl J., Selenko N.,
RA Prager E., Staffler G., Madic O., Stockinger H., Knapp W.;
RT "Identification of human CD93 as the phagocytic C1q receptor (C1qR)
RT by expression cloning.";
RL J. Leukoc. Biol. 71:133-140(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavridis G., Almeida J.P., Babbage A.K., Bagguley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P., Clee C.M.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Cunn M.,
RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A.G., Coville G.J., Deadman R., Dhali P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levenslaio M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A.C., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
RN [4]

```

```

RP SEQUENCE FROM N.A.
RC TISSUE=Leukocyte;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [5]
RP CHARACTERIZATION.
RX PubMed=11994479;
RA McCreel E.P., Iwakaki N., Akatsu H., Morgan B.P., Gasque P.;
RT "Human C1qR is identical with CD93 and the MNI-11 antigen but does
RT not bind C1q.";
RL J. Immunol. 168:5222-5232(2002).
RN [6]
RP O-GLYCOSYLATION.
RX PubMed=10092817;
RA Nepomuceno R.R., Ruiz S., Park M., Tenner A.J.;
RT "C1qR is a heavily O-glycosylated cell surface protein involved in
RT the regulation of phagocytic activity.";
RL J. Immunol. 162:3583-3589(1999).
CC -!- FUNCTION: Receptor (or element of a larger receptor complex) for
CC C1q, mannose-binding lectin (MBL2) and pulmonary surfactant
CC protein A (SPA). May mediate the enhancement of phagocytosis in
CC monocytes and macrophages upon interaction with soluble defense
CC collagens. May play a role in intercellular adhesion.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Highly expressed in endothelial cells,
CC platelets, cells of myeloid origin, such as monocytes and
CC neutrophils. Not expressed in cells of lymphoid origin.
CC -!- PTM: N- and O-glycosylated.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.
CC -!- CAUTION: Has been sometimes referred to as a collectin receptor.
CC -!- CAUTION: According to Ref.5, C1q is not a ligand for C1QR1.
CC -!- DATABASE: NAME=PROW; NOTE=PROW 3:1-6(2001); /467246456.g.htm".
CC WWW="http://www.ncbi.nlm.nih.gov/prov/guide/467246456.g.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U94333; AAB53110.1; -.
DR EMBL: AL118508; CAC00597.1; -.
DR EMBL: BC028075; AAB28075.1; -.
DR HSSP: P35555; IEMN.
DR MIM: 120577; -.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR001304; Lectin_C.
DR InterPro: IPR001187; Tissue_factor.
DR Pfam: PF00008; EGF; 5.
DR Pfam: PF00059; lectin_c; 1.
DR Pfam: PF01108; Tissue_fac; 1.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00179; EGF_CA; 3.
DR SMART: SM00001; EGF_like; 2.
DR PROSITE: PS00010; ASX_HYDROXYL; 3.
DR PROSITE: PS00615; C_TYPE_LECTIN_1; FALSE_NEG.
DR PROSITE: PS00441; C_TYPE_LECTIN_2; 1.
DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS01187; EGF_CA; 3.
KW Receptor; EGF-like domain; Signal; Transmembrane; Glycoprotein;
KW Repeat; Lectin; Polymorphism.
FT SIGNAL 1
FT CHAIN 21
FT DOMAIN 22 652
FT DOMAIN 24 580
FT TRANSMEM 591 601
FT TRANSMEM 602 652
FT DOMAIN 602 174
FT DOMAIN 260 301

```

FT CARBOHYD 494 494 O-LINKED (GLYCOSAMINOGLYCAN) (BY  
FT SIMILARITY).  
SQ SEQUENCE 577 AA; 61867 MW; B20E50B0FE745014 CRC64;  
Query Match 65.1%; Score 1899; DB 1; Length 577;  
Best Local Similarity 65.5%; Pred. No. 5.2e-118;  
Matches 341; Conservative 46; Mismatches 126; Indels 8; Gaps 4;  
Qy 1 MLGVVLGALAGLGFAPAEPQGSQCVHDCALYVGPATFELNASQICDLGRHLM 60  
Db 1 MLGIFLVLPASLGLSALAKLQPTGSCQVHECFALFGPATFELNASQICDLGRHLM 60  
Qy 61 TVRSSVAADVISLLSQQSMDLGP--WIGLQLPQCGDDPVHLGLRGLFGQWVTGDNNTSY 119  
Db 61 TVRSSVAADVISLLSQQSMDLGP--WIGLQLPQCGDDPVHLGLRGLFGQWVTGDNNTSY 117  
Qy 120 SRWRLDNLGAPLCPGVAYSAEATVPSEPIWEQOCEVKADGFLCEPHFPATCRPLA 179  
Db 118 SRWARPNDQAPLCPGVAYSAEATVPSEPIWEQOCEVKADGFLCEPHFPATCRPLA 177  
Qy 180 VEP-GAATAAVSYTYGTPFAARGADQALPVGSSAAVAPLGLQLMCTAPPQVQGHWARE 238  
Db 178 VNTPRPAHISSTYNTYTPFGVSGADQTLPLVGSAAVEPLGLVLCRAPPTSEGHWARE 237  
Qy 239 APGAWDCSVNGGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDCHEFCVP 298  
Db 238 ATGAWNCSVNGGCEYLNCSTNEPRCLCPRDMDLQADGRSCARPVYVQSCNELCEHFCVS 297  
Qy 299 NPDPGYSYCMCTGYRLAQRHCEVDVDCILEPSPCPORCVNTGGFECHECHYPNYDLV 358  
Db 298 NAEVPGYSYCMCTGYRLAQRHCEVDVDCILEPSPCPORCVNTGGFECHECHYPNYDLV 357  
Qy 359 DGEVCEVDPDFCFRANCYQCPNLTQSYLVCVCAEGFAPIPHEPRQCMFCNQACPADCD 418  
Db 358 DGEVCEVDPDFCFRANCYQCPNLTQSYLVCVCAEGFAPIPHEPRQCMFCNQACPADCD 417  
Qy 419 PNTQASCEPEGYLIDDFICTIDECENGFCGVCNHLNPGTFECICGPDSPDLSRHIGT 478  
Db 418 PNSTVCEPEGYLIDDFICTIDECENGFCGVCNHLNPGTFECICGPDSPDLSRHIGT 477  
Qy 479 DCDSEKV---DGGDSGSGEPPTPTGTLTPPAVGLVHSG 516  
Db 478 DCDPIPVREDTKEBEGSGEPPTPTGTLTPPAVGLVHSG 518

RESULT 3

TRBM\_BOVIN  
ID TRBM\_BOVIN STANDARD; PRT; 356 AA.  
AC P06579;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Thrombomodulin (Fetomodulin) (TW) (Fragment).  
GN THBD.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87067408; PubMed=3024152;  
RA Jackman R.W., Beeler D.L., Vandewater L., Rosenberg R.D.;  
RT "Characterization of a thrombomodulin cDNA reveals structural  
RL similarity to the low density lipoprotein receptor."  
RL Proc. Natl. Acad. Sci. U.S.A. 83:8834-8838(1986).  
CC -1- FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR  
CC THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS  
CC COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE  
CC ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA  
CC SCISSORS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,  
CC FACTOR VA AND FACTOR VIIIa, AND THEREBY REDUCES THE AMOUNT OF  
CC THROMBIN GENERATED.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING  
CC THROMBOMODULIN.  
CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; M14657; AAA30785.1; -.  
CC PIR; A25918; A25918.  
CC HSP; P07204; 1TMR.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_Ca.  
DR Pfam; PF00008; EGF; 5.  
DR SMART; SM00179; EGF\_Ca; 1.  
DR SMART; SM00001; EGF-like; 3.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 2.  
DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
DR PROSITE; PS01186; EGF\_2; 3.  
DR PROSITE; PS01187; EGF\_Ca; 2.  
KW Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane;  
KW Glycoprotein; EGF-like domain.  
FT NON\_TER 1 1  
FT DOMAIN 1 296 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 297 320 POTENTIAL.  
FT DOMAIN 321 356 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 17 57 EGF-LIKE 1.  
FT DOMAIN 60 98 EGF-LIKE 2.  
FT DOMAIN 99 137 EGF-LIKE 3. CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 139 179 EGF-LIKE 4.  
FT DOMAIN 178 214 EGF-LIKE 5.  
FT DOMAIN 215 254 EGF-LIKE 6. CALCIUM-BINDING (POTENTIAL).  
FT DISULFID 21 32 BY SIMILARITY.  
FT DISULFID 28 41 BY SIMILARITY.  
FT DISULFID 43 56 BY SIMILARITY.  
FT DISULFID 64 72 BY SIMILARITY.  
FT DISULFID 68 82 BY SIMILARITY.  
FT DISULFID 84 97 BY SIMILARITY.  
FT DISULFID 103 114 BY SIMILARITY.  
FT DISULFID 110 123 BY SIMILARITY.  
FT DISULFID 125 136 BY SIMILARITY.  
FT DISULFID 143 152 BY SIMILARITY.  
FT DISULFID 148 162 BY SIMILARITY.  
FT DISULFID 164 178 BY SIMILARITY.  
FT DISULFID 182 191 BY SIMILARITY.  
FT DISULFID 187 199 BY SIMILARITY.  
FT DISULFID 201 213 BY SIMILARITY.  
FT DISULFID 219 228 BY SIMILARITY.  
FT DISULFID 224 237 BY SIMILARITY.  
FT DISULFID 239 253 BY SIMILARITY.  
FT CARBOHYD 271 271 O-LINKED (GLYCOSAMINOGLYCAN) (BY  
FT SIMILARITY).  
SQ SEQUENCE 356 AA; 37795 MW; 29B41F097ABE1093 CRC64;  
Query Match 37.5%; Score 1092.5; DB 1; Length 356;  
Best Local Similarity 63.4%; Pred. No. 3.5e-65;  
Matches 189; Conservative 30; Mismatches 64; Indels 15; Gaps 6;  
Qy 229 GAVQGHWAREAPGAWDCSVNGGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSC 288  
Db 5 GTEGRWSREAPACWGVGRCQHECKGSAGASNCCLCPADAALQADGRSCGLPAEHPC 64  
Qy 289 NDLCEHFCVPNPDPGYSYCMCTGYRLAQRHCEVDVDCILEPSPCPORCVNTGGFEC 348  
Db 65 HQJCEHFC--HLHGLGNLYTCICAGYLAADQRHCEVDVDCILEPSPCPORCVNTGGFEC 122  
Qy 349 CHCYPNYDLVDCGCEVDFVDFCFRANCYQCPNLTQSYLVCVCAEGFAPIPHEPRQCMFC 408

CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC	CC
----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----

RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.K., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Levaslaiho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConachie L.J., Mclay K., McMurray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.F., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20";  
RL Nature 414:865-871(2001).  
RN [6]  
RP CARBOHYDRATE-LINKAGE SITE SER-492, AND MUTAGENESIS.  
RX MEDLINE-94029900; PubMed-8216207;  
RA Grinnell B.W.;  
RT "Identification of the predominant glycosaminoglycan-attachment site  
in soluble recombinant human thrombomodulin: potential regulation of  
functionality by glycosyltransferase competition for serine474.";  
RL Biochem. J. 295:131-140(1993).  
RN [7]  
RP STRUCTURE BY NMR OF 389-407.  
RX MEDLINE-96007474; PubMed-7559494;  
RA Adler M., Seto M.H., Nitecki D.E., Lin J.H., Light D.R., Morser J.;  
RT "The structure of a 19-residue fragment from the C-loop of the fourth  
epidermal growth factor-like domain of thrombomodulin.";  
RL J. Biol. Chem. 270:23366-23372(1995).  
RN [8]  
RP STRUCTURE BY NMR OF 364-407.  
RX MEDLINE-96100636; PubMed-8528067;  
RA Meininger D.P., Hunter M.J., Komives E.A.;  
RT "Synthesis, activity, and preliminary structure of the fourth  
EGF-like domain of thrombomodulin.";  
RL Protein Sci. 4:1683-1695(1995).  
RN [9]  
RP STRUCTURE BY NMR OF 427-444.  
RX MEDLINE-95034791; PubMed-7947766;  
RA Srinivasan J., Hu S., Hrabal R., Zhu Y., Komives E.A., Ni F.;  
RT "Thrombin-bound structure of an EGF subdomain from human  
thrombomodulin determined by transferred nuclear Overhauser  
effects.";  
RL Biochemistry 33:13553-13560(1994).  
RN [10]  
RP STRUCTURE BY NMR OF 427-444.  
RX MEDLINE-96276211; PubMed-8745396;  
RA Hrabal R., Komives E.A., Ni F.;  
RT "Structural resiliency of an EGF-like subdomain bound to its target  
protein, thrombin.";  
RL Protein Sci. 5:195-203(1996).  
RN [11]  
RP STRUCTURE BY NMR OF 405-444.  
RX MEDLINE-98035729; PubMed-9367781;  
RA Sampoli Benitez B.A., Hunter M.J., Meininger D.P., Komives E.A.;  
RT "Structure of the fifth EGF-like domain of thrombomodulin: an  
EGF-like domain with a novel disulfide-bonding pattern.";  
RL J. Mol. Biol. 273:913-926(1997).  
RN [12]  
RP VARIANT TED TYR-486.  
RX MEDLINE-95111115; PubMed-7811989;  
RA Oehlin A.-K., Marlar R.A.;  
RT "The first mutation identified in the thrombomodulin gene in a  
45-year-old man presenting with thromboembolic disease.";  
RL Blood 85:330-336(1995).  
RN [13]  
RP VARIANT TED Y-486, AND VARIANTS T-43; A-79; S-495 AND L-501.

RX MEDLINE-97341986; PubMed-9198186;  
RA Oehlin A.-K., Norlund L., Marlar R.A.;  
RT "Thrombomodulin gene variations and thromboembolic disease.";  
RL Thromb. Haemost. 78:396-400(1997).  
RN [14]  
RP VARIANT VAL-473.  
RX MEDLINE-9720518; PubMed-9157575;  
RA Norlund L., Holm J., Zoller B., Oehlin A.-K.;  
RT "A common thrombomodulin amino acid dimorphism is associated with  
myocardial infarction.";  
RL Thromb. Haemost. 77:248-251(1997).  
RN [15]  
RP VARIANT THR-43.  
RX MEDLINE-99057299; PubMed-9843165;  
RA Doggen C.J.M., Kunz G., Rosendaal F.R., Lane D.A., Vos H.L.,  
RA Stubbs P.J., Manger Cats V., Ireland H.;  
RT "A mutation in the thrombomodulin gene, 127G to A coding for Ala25Thr,  
and the risk of myocardial infarction in men.";  
RL Thromb. Haemost. 80:743-748(1998).  
RN [16]  
RP VARIANT VAL-473.  
RX MEDLINE-21143723; PubMed-11245641;  
RA Wu K.K., Aleksic N., Ann C., Boerwinkle E., Folsom A.R.,  
RA Juneja H.;  
RT "Thrombomodulin Ala455Val polymorphism and risk of coronary heart  
disease.";  
RL Circulation 103:1386-1389(2001).  
CC [-] FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR  
THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS  
COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE  
ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA  
SCISSONS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,  
FACTOR VA AND FACTOR VIIIA, AND THEREBY REDUCES THE AMOUNT OF  
THROMBIN GENERATED.  
CC [-] SUBCELLULAR LOCATION: Type I membrane protein.  
CC [-] TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING  
THROMBOMODULIN.  
CC [-] POLYMORPHISM: VARIATIONS IN THBD ARE ASSOCIATED WITH AN INCREASED  
RISK OF DEVELOPING THROMBOEMBOLIC DISEASES (TED).  
CC [-] DISEASE: DEFECTS IN THBD COULD BE THE CAUSE OF INHERITED TED,  
ALSO CALLED INHERITED THROMBOPHILIA. PATIENTS WITH TED HAVE  
DEFECTS OF THE HAEMOPOIETIC SYSTEM WHICH CREATES A TENDENCY TO THE  
OCCURRENCE OF THROMBOSIS. TED PLAY AN IMPORTANT ROLE IN THE  
PATHOGENESIS OF VARIOUS CARDIOVASCULAR DISORDERS.  
CC [-] SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.  
CC [-] DATABASE: NAME=PROW; NOTE=cd guide cd141 entry;  
WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd141.htm".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X05495; CAA29045.1; -  
DR EMBL: M16552; AAB59508.1; -  
DR EMBL: J02973; AAA61175.1; -  
DR EMBL: D00210; BAA00149.1; -  
DR EMBL: AL049651; CAB51954.1; -  
DR PIR: A27073; A27073.  
DR PIR: A28307; A28307.  
DR PIR: A29680; A29680.  
DR PDB: 1EGT; 15-NOV-95.  
DR PDB: 1FGD; 20-JUN-96.  
DR PDB: 1FGE; 20-JUN-96.  
DR PDB: 1TMR; 08-JUN-95.  
DR PDB: 1ZAO; 29-JAN-96.  
DR PDB: 1ADX; 24-DEC-97.  
DR PDB: 2ADX; 24-DEC-97.  
DR GlycoSuiteDB; P07204; -  
DR Genew; HGNC:11784; THBD.

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:20:02 ; Search time 10.5 Seconds  
(without alignments)  
2038.265 Million cell updates/sec

Title: US-09-509-994-1  
Perfect score: 2916  
Sequence: 1 MLGVVLGALAGLPAP.....PSPTPGSTLTPPAVLVHSG 516

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2912	99.9	575	1 TRBM_HUMAN	P07204 homo sapien
2	1899	65.1	577	1 TRBM_MOUSE	P15306 mus musculus
3	1092.5	37.5	356	1 TRBM_BOVIN	P08579 bos taurus
4	583	20.0	652	1 CD93_HUMAN	Q9npj3 homo sapien
5	543	18.6	644	1 CD93_MOUSE	O89103 mus musculus
6	530.5	18.2	643	1 CD93_RAT	Q9et61 rattus norv
7	373	12.8	2907	1 FN2_MOUSE	G61555 mus musculus
8	367.5	12.6	1184	1 FN2_HUMAN	P98095 homo sapien
9	367	12.6	2871	1 FN1_BOVIN	P98133 bos taurus
10	366	12.6	2871	1 FN1_PIG	Q9tv36 sus scrofa
11	365	12.5	2911	1 FN2_HUMAN	P35556 homo sapien
12	363	12.4	2871	1 FN1_HUMAN	P35555 homo sapien
13	361	12.4	2871	1 FN1_MOUSE	G61554 mus musculus
14	358	12.3	956	1 MN2_HUMAN	O00339 homo sapien
15	357	12.2	1221	1 FN2_MOUSE	P37889 mus musculus
16	329.5	11.3	1964	1 NTC4_MOUSE	P31695 mus musculus
17	328.5	11.3	956	1 MN2_MOUSE	O08746 mus musculus
18	327.5	11.2	1712	1 FN1_RAT	Q00918 rattus norv
19	324.5	11.1	712	1 FN1_CAEEL	O77469 caenorhabdi
20	320.5	11.0	2321	1 NTC3_HUMAN	Q9um47 homo sapien
21	320	11.0	1394	1 LTB3_HUMAN	P22064 homo sapien
22	320	11.0	1595	1 LTB4_HUMAN	Q14766 homo sapien
23	315.5	10.8	443	1 FN4_HUMAN	O95967 homo sapien
24	313	10.7	703	1 FN1_HUMAN	P23142 homo sapien
25	308.5	10.6	443	1 FN4_CRIGR	O55058 cricetus
26	308.5	10.6	2003	1 NTC4_HUMAN	Q99466 homo sapien
27	308	10.6	448	1 FN5_HUMAN	O9ubx5 homo sapien
28	308	10.6	705	1 FN1_MOUSE	Q08879 mus musculus
29	307	10.5	684	1 FN1_CHICK	O73775 gallus gall
30	306.5	10.5	443	1 FN4_MOUSE	Q9wv19 mus musculus
31	302	10.4	448	1 FN5_MOUSE	O9wv19 mus musculus
32	296	10.2	448	1 FN5_RAT	Q9wv18 rattus norv
33	296	10.2	2531	1 NTC1_MOUSE	Q01705 mus musculus

34	293.5	10.1	2319	1 NTC3_RAT	Q9rl72 rattus norv
35	291	10.0	2470	1 NTC2_MOUSE	O35516 mus musculus
36	288.5	9.9	493	1 FN3_HUMAN	Q12805 homo sapien
37	286	9.8	2471	1 NTC2_RAT	Q9gw30 rattus norv
38	284	9.7	2471	1 NTC2_HUMAN	Q04721 homo sapien
39	283.5	9.7	2437	1 NTC1_BRARE	P46530 brachydanio
40	280.5	9.6	2703	1 NTC3_DROME	P07207 drosophila
41	280	9.6	2318	1 NTC3_MOUSE	Q61982 mus musculus
42	277	9.5	2531	1 NTC1_RAT	Q07008 rattus norv
43	276.5	9.5	493	1 FN3_RAT	O35568 rattus norv
44	271	9.3	2556	1 NTC1_HUMAN	P46531 homo sapien
45	269.5	9.2	1217	1 EGF_MOUSE	P01132 mus musculus

ALIGNMENTS

RESULT 1

ID	TRBM_HUMAN	STANDARD;	PRT;	575 AA.
AC	P07204;			
DT	01-APR-1988 (Rel. 07, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Thrombomodulin precursor (Fetomodulin) (TM) (CD141 antigen).			
GN	THBD OR THRM.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88004395; PubMed=2820710;			
RA	SUZUKI K., Kusumoto H., Dayashiki Y., Nishioka J., Maruyama I.,			
RA	Zushi M., Kawahara S., Honda G., Yamamoto S., Horiguchi S.;			
RT	"Structure and expression of human thrombomodulin, a thrombin			
RT	receptor on endothelium acting as a cofactor for protein C			
RT	activation."			
RL	EMBO J. 6:1891-1897(1987).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88024950; PubMed=2822087;			
RA	Wen D., Dittman W.A., Ye R.D., Deaven L.L., Majerus P.W., Sadler J.E.;			
RA	"Human thrombomodulin: complete cDNA sequence and chromosome			
RT	localization of the gene."			
RL	Biochemistry 26:4350-4357(1987).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87317665; PubMed=2819876;			
RA	Jackman R.W., Beeler D.L., Fritze L., Soff G., Rosenberg R.D.;			
RT	"Human thrombomodulin gene is intron depleted: nucleic acid sequences			
RT	of the cDNA and gene predict protein structure and suggest sites of			
RT	regulatory control."			
RL	Proc. Natl. Acad. Sci. U.S.A. 84:6425-6429(1987).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88227901; PubMed=2836377;			
RA	Shirai T., Shiojiri S., Ito H., Yamamoto S., Kusumoto H.,			
RA	Dayashiki Y., Maruyama I., Suzuki K.;			
RT	"Gene structure of human thrombomodulin, a cofactor for thrombin-			
RL	catalyzed activation of protein C."			
RL	J. Biochem. 103:281-285(1988).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21638749; PubMed=11780052;			
RA	Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,			
RA	Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,			
RA	Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,			
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,			
RA	Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,			
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,			
RA	Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,			
RA	Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,			

Search completed: December 16, 2002, 17:26:48  
Job time : 25.5 secs

R:Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000

A:Reference number: Z23035

A:Accession: T46488

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-741 <AAA>

A:Cross-references: EMBL:AL137638

A:Experimental source: adult testis; clone DKFz434J065

C:Genetics:

A:Note: DKFz434J065.1

Query Match 12.3%; Score 359; DB 2; Length 741;  
Best Local Similarity 31.6%; Pred. No. 1.5e-14;  
Matches 86; Conservative 41; Mismatches 101; Indels 44; Gaps 15;

Qy 245 CSVENGGCEHACNAIPGAPRCQCPAGAAQADGRSCTA-SATQSCNDLCEHFCVNPDPQ 303

Db 68 CAMEDHNCQELCVNVPSCVQCVGYALAEADGKRCVAVDYCASENHGCEHCNAD-- 124

Qy 304 GSYSCMETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHECHYNYDL-VDGEC 362

Db 125 GSYLCQHEGFPALNPEKTKTKIDYCASNNHGCQHECVNTDDSYSCHLAGFTINPDKKT 184

Qy 363 VEPVDPGF--RANCEYOCQPLNQTSLVCVCAEGFAPIPH-----EPHRCQMF 408

Db 185 CRRINYCALNKPGEHECVNNEE--SYICRHRGTYLDPNKTCSRVDHCAQQDHGCEQLC 243

Qy 409 NQTACPADCPDNTQAS--CECPGYILDDGF-ICTDIDEC---ENGFCSCVGNHLPGFT 462

Db 244 -----LNTEDSFVCCSEGLNEDLKTCSRVDYCLLSDHG--CEYCVNMDRSF 291

Qy 463 ECIGPDSALVRHGTDCDSKVDG---GDSG 491

Db 292 ACQC-PEGHVLRSDGKTC--AKLDSALGDHG 320

#### RESULT 13

A49457

fibulin-2 precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 02-Aug-2002

C:Accession: A49457; S74095

R:Pan, T.C.; Sasaki, T.; Zhang, R.Z.; Faessler, R.; Timpl, R.; Chu, M.L.

J. Cell Biol. 123, 1269-1277, 1993

A>Title: Structure and expression of fibulin-2, a novel extracellular matrix protein with

A:Reference number: A49457; MUID:94064787; PMID:8245130

A:Accession: A49457

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1221 <PAN>

A:Cross-references: GB:X75285; NID:g437046; PIDN:CAA53040.1; PID:g437047

R:Sasaki, T.; Mann, K.; Murphy, G.; Chu, M.L.; Timpl, R.

Eur. J. Biochem. 240, 427-434, 1996

A>Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix met

A:Reference number: S74094; MUID:96439073; PMID:8841408

A:Accession: S74095

A:Molecule type: protein

A:Residues: 236-238, 'X', 240-247; 260-275; 336-344, 'L', 346-361; 405-426; 566-568, 'EM', 569-589

C:Superfamily: fibulin-2; EGF homology

C:Keywords: calcium binding; duplication; extracellular matrix; glycoprotein; homotrimer

F;942-978/Domain: EGF homology <EGF>

Query Match 12.2%; Score 357; DB 2; Length 1221;  
Best Local Similarity 29.5%; Pred. No. 3e-14;  
Matches 123; Conservative 39; Mismatches 141; Indels 114; Gaps 27;

Qy 154 EEQCEVKAD-GFLCEHFHATC-----RPLAV-----EPGAAAVS-----IT 192

Db 521 EGQSCSNPNIGYPCN-HVMLSCGEGEPLIVPEVRPPPEFAAPRVSEMASREALS 579

Qy 193 YGT----PFAARGADFO---ALP-----VGSSAAVAPLGLQLM-----CTAPP 229

Db 580 LGTEAEPLNSLPDQDDQDECLMLPGLCQHLINTVGSYRCACFPPELODGRTPDRG 639

Qy 230 AVOGHWARE-APGAWDCSV-----ENGGEHACNAIPGAPRCQCPAGAA 272

Db 640 APOLDTARESAAPRSESAQVSPNTPPLVPQPNCTCKDNGPCRCQVCRVVGDTAMCSCFPGYA 699

Qy 273 LQADGRSC-----TASATQSCNDLCEHFCVNPDPQSGSYSC-----MCETGYRLAADQHR 322

Db 700 IMADGVSCEDQDECLMGTHDCS--WKQFCV---NTLGSFYCVNHTVLCAGYILNA-HRK 753

Qy 323 CEDVDDCILEPSPC--PQRCVNTQGGFECH---CYPNTDLDVDCGEVPEVDPFCFRA--NC 374

Db 754 CVDINECVTLHTCTRAEHCNVNTPGFSQCYKALTCFPGVILDTGECTD-VDECVTGTNC 812

Qy 375 E-YQCOPLNQTSLYCV---CAEGFAPIPH-----EPHRCQMFNQTACPA 415

Db 813 QAGFSCQN-TKGSFYCAQRQCMDFLQDPEGNCVDINECTSLLEPCRSFGSCINTVGSY 871

Qy 416 DCDPNTQASCECPGY-ILDDGFICTDIDCEGNGFCSG--VCHNLPGTFECICGP 468

Db 872 TCORNPLV---CGRGYHANESEGVVDVNECETGVHRCGEGQLCYNLPGSRDCKP 925

#### RESULT 14

T43210

fibulin-1D precursor - Caenorhabditis elegans (fragment)

C:Species: Caenorhabditis elegans

C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 02-Sep-2000

C:Accession: T43210

R:Barth, J.L.; Argraves, K.M.; Roark, E.F.; Little, C.D.; Argraves, W.S.

submitted to the EMBL Data Library, June 1998

A>Description: Identification of chicken and C. elegans fibulin-1 homologs and charac

A:Reference number: Z22337

A:Accession: T43210

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-589 <BAR>

A:Cross-references: EMBL:AF070477; PIDN:AAC24035.1

C:Genetics:

A:Note: intron positions not resolved (incomplete sequence)

C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 11.4%; Score 331; DB 2; Length 589;

Best Local Similarity 28.1%; Pred. No. 5.9e-13;

Matches 126; Conservative 49; Mismatches 166; Indels 108; Gaps 28;

Qy 103 GPLRGF-QWVTGDNNTSYRWARLDLNGAPLCGLCAVSAEAATVPSEPI-----WEQ 156

Db 19 GCLRSFNKCCNGDIEITH---ASEIITGRPLNDPHVLHGLDRCASSKCHLCHDRGKRV 75

Qy 157 QCEVKA-----DGFLCEHFHP--ATCRPLAVE-----PGAAAAVAVITYGTTPA--- 198

Db 76 ECSCRSGFDLADPGMACVDHIDECATLDDCLESQRCLNTPGSKCIRTLSGCTGYAMDS 135

Qy 199 --ARGADFOALPVGSSAAVAPLGLQMLCTAPCAVQ-----CHWAREAPGAWDCS---- 246

Db 136 ETERCRDVEDCNLGSN-----DCGLPYQCRNTQGSYRCDAKCKGDLQELNPGTGETSITC 191

Qy 247 -----VENGGCE-----HACNA-----IPGAPRCQ-----CPAGAAQADGRSCTA 282

Db 192 PNGYYPKNGMNDIDSCVTGHNGCGAGEECVNTPGSRFCQKGNLCAHGEVNV----- 243

Qy 283 SATQSCNDL--CEH-----FCVPNPDPGYSYSCMETGYRLAADQHRCEVDVDCIL---- 331

Db 244 GATGFCDVNECCQGVGSGMECI---NLPGYTKCKGPGYEFNDAKKRCEVDDECIFKAG 300

Qy 332 EPSPCQRCVNTQGGFECHECHYNYDLV-DGECVEPVPDPCR--ANCEYOCQPLNQTSLYLC 388

Db 301 HVCDSLAEICNTIGSGCEKCKPGQLASDGRRCEDVNECTGTAAECQCVNI-PGSYQC 359

Qy 389 VCAEGFAPIP-----EPHRCQMF--NOTACPADCPDNTQAS--CECPEGY-ILDDGFI 438

Db 360 ICDRGFALGDPGPKCEDIDECISWAGSNDLNGGC--INTKGYLCQCPGKYIQDPGR 418



R:Perreira, L.V.; D'Alessio, M.; Ramirez, F.; Lynch, J.; Sykes, B.; Pangillan, T.; Bonad  
Hum. Mol. Genet. 2, 961-968, 1993  
A:Title: Genomic organization of the sequence coding for fibrillin, the defective gene B  
A:Reference number: I54355; MUID:93372860; PMID:8364578  
A:Accession: I54355  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 132-3002 <PER>  
A:Cross-references: GB:13923; NID:9306745; PIDN:AAB02036.1; PID:g306746  
R:Maslen, C.L.; Corson, G.M.; Maddox, B.K.; Glanville, R.W.; Sakai, L.Y.  
Nature 352, 334-337, 1991  
A:Title: Partial sequence of a candidate gene for the Marfan syndrome.  
A:Reference number: S17064; MUID:91304568; PMID:1852207  
A:Accession: S17064  
A:Molecule type: mRNA  
A:Residues: 1030-3002 <MAS>  
A:Cross-references: EMBL:X63556  
R:Dietz, H.C.; Valle, D.; Francomano, C.A.; Kendzior, R.J.  
Science 259, 680-683, 1993  
A:Title: The skipping of constitutive exons in vivo induced by nonsense mutations.  
A:Reference number: I59574; MUID:93157831; PMID:8430317  
A:Accession: I59574  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 2217-2288, 'I', 2290-2325 <RES>  
A:Cross-references: GB:S54426; NID:9264860; PIDN:AAB25244.1; PID:g264861  
R:Lee, B.; Godfrey, M.; Vitale, E.; Horl, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras, P.  
Nature 352, 330-334, 1991  
A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two differe  
A:Reference number: S17062; MUID:91304567; PMID:1852206  
A:Accession: S17062  
A:Molecule type: mRNA  
A:Residues: 'VLTVTFVFIYSYNKML', 944-1444 <LEE1>  
A:Cross-references: EMBL:X62008; NID:g91398; PIDN:CAB56534.1; PID:g5924015  
A:Accession: S62111  
A:Molecule type: protein  
A:Residues: 1166-1176, 'X', 1178-1180, 'D', 1182-1185 <LEE2>  
R:Maddox, B.K.; Sakai, L.Y.; Keene, D.R.; Glanville, R.W.  
J. Biol. Chem. 264, 21381-21385, 1989  
A:Title: Connective tissue microfibrils. Isolation and characterization of three large F  
A:Reference number: A34198; MUID:90078246; PMID:2512293  
A:Accession: A34198  
A:Molecule type: protein  
A:Residues: 565-575, 1890-1892, 'I', 1894-1900 <MAD>  
C:Comment: Fibrillin is a major component of elastin-associated microfibrils.  
C:Genetics:  
A:Gene: GDB:FBN1  
A:Cross-references: GDB:127115; OMIM:134797; OMIM:154700  
A:Map position: 15q21.1-15q21.1  
A:Introns: 2236/1; 2258/1; 2297/1  
C:Superfamily: fibrillin 1; EGF homology  
C:Keywords: alternative splicing; calcium binding; extracellular matrix; glycoprotein; M  
F:1-3002/Product: fibrillin (5'-region exon A splice form) (fragment) #status predicted  
F:1332-1367/Domain: fibrillin (5'-region exon C splice form) #status predicted <MATC>  
F:1457-1492/Domain: EGF homology <EGF>  
F:2262-2295/Domain: EGF homology <EGF>  
F:2262-2295/Domain: EGF homology <EGF1>  
Query Match 12.4%; Score 363; DB 2; Length 3002;  
Best Local Similarity 30.8%; Pred. No. 2.8e-14;  
Matches 111; Conservative 22; Mismatches 105; Indels 122; Gaps 20;  
Qy 244 DCSYVNGGCEHACNAIFGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP- 300  
Db 1331 ECSIMNGGCTCTNSGSEYSCSQPGFALMPDQSRCT-----DIDE--CEDNPNIC 1380  
Qy 301 -----DQPGYSQCMCEGTVLAADQHRCEVDVDCILEPSPC-PQRCVNTQGGFECCHY 353  
Db 1381 DGGQCTNIPGEYRCLVDGFWASDMKTCVDVNECDLNPICLSGTCENTKGSFICHDM 1440  
Qy 354 NYDLVDGE--CPEVPDPC--FRANCEYQCPQLNQT-SYLCVCAEGFA-----PIP 398  
Db 1441 GYSKKKGTGCTD-INECEIGAHNCGKHAVCTNTAGSFKCSQSPGWIGDGKICTDLDECS 1499

Qy 399 HEPHRCQMFQNTACPADCDPNTQAS--CECEGYLDGDFICTDIDEC-EN----- 447  
Db 1500 NTHMCSQH-----ADC-KNTMGSYRCLCKEGY-TGDGFTCTDLDECSNLCGNGQ 1550  
Qy 448 -----GGF-----CS-----GVCHNLPGTFECICGPDLSALV 473  
Db 1551 CLNAPGGYRCECDMGFVPSADGKACEDIDECSLPNICVFGTCHNLPGFRCCEIGEYELD 1610  
Qy 474 RHIG-----TDCDSKGVGDGSGSGEPSPPTGS-----TLTPPAVGLV 513  
Db 1611 RSGGCTDVNECLDPTTCISGCVN-----TPGSYICDPPDFELNPTRVGCV 1658  
RESULT 11  
A55624  
fibrillin-1 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change 02-Aug-2002  
C:Accession: A55624  
R:Yin, W.; Smiley, E.; Germiller, J.; Sanguinetti, C.; Lawton, T.; Perreira, L.; Ramire  
J. Biol. Chem. 270, 1798-1806, 1995  
A:Title: Primary structure and developmental expression of Fbn-1, the mouse fibrillin  
A:Reference number: A55624; MUID:95130561; PMID:7829516  
A:Accession: A55624  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-2871 <YIN>  
A:Cross-references: GB:L29454; NID:g575509; PIDN:AAA56840.1; PID:g575510  
C:Genetics:  
A:Gene: Fbn-1  
C:Superfamily: fibrillin 1; EGF homology  
F:1201-1236/Domain: EGF homology <EGF>  
Query Match 12.4%; Score 361; DB 2; Length 2871;  
Best Local Similarity 29.4%; Pred. No. 3.5e-14;  
Matches 117; Conservative 26; Mismatches 127; Indels 138; Gaps 21;  
Qy 212 SAAVAPLGL-----QLMCTAPGAVOGHWAREAGWDSCVSENGCEHACNAIFGAPRC 265  
Db 1162 SANLCPHGRVNLIGKYQACNPGYHPTDRLFCVDIDECSIMNGGCTCTNSDGSYEC 1221  
Qy 266 QCPAGAAQADGRSCTASATQSCNDLCEHFCVNP-----DQPGYSQCMCEGTVRL 316  
Db 1222 SQPGFALMPDQSRCT-----DIDQ--CEDNPNICDGGQCTNIPGEYRCLVDGFWA 1271  
Qy 317 AADQHRCEVDVDCILEPSPC-PQRCVNTQGGFECCHYPNYDLVDGE--CPEVPDPC--FR 371  
Db 1272 SEMDKTCVDVNECDLNPICLSGTCENTKGSFICHDMGYSGKKGKTGCTD-INECEIGA 1330  
Qy 372 ANCEYQCPQLNQT-SYLCVCAEGFA-----PIPEHRCQMFQNTACPADCDPN 420  
Db 1331 HNGRHAVCTNTAGSFKCSQSPGWIGDGKICTDLDECSNGTHMCSQH-----ADC-KN 1382  
Qy 421 TOAS--CECEGYLDGDFICTDIDEC-EN-----GGF----- 450  
Db 1383 TMGSYRCLCKDGY-TGDGFTCTDLDECSNLCGNGQCLNAPGGYRCECDMGFVPSADG 1441  
Qy 451 -----CS-----GVCHNLPGTFECICGPDLSALVRHIG-----TDCDSK 484  
Db 1442 KACEDIDECSLPNICVFGTCHNLPGFRCCEIGEYELDRSGGNTDVNECLDPTTCISGN 1501  
Qy 485 VDGDSGSGEPSPPTGS-----TLTPPAVGLV 513  
Db 1502 CVN-----TPGSYTCDCSPDFELNPTRVGCV 1527  
RESULT 12  
T46488  
hypothetical protein DKFp434J065.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
C:Accession: T46488

Db 683 CKDNGPCKQVC---STVGGSAICSPPGYALMADGVSCEDINECVTDLHTCSRGHCNVT 739

Qy 344 QGGFECH---CYPNDYLDVGCVEPVPDPCFRANCEYQCQPLNQTSYLCV----- 389

Db 740 LGSFHCYKALTCEPGYALKDGC-EDVDEC--AMGTHTCQP-----GFLQCNKSGSYCOA 792

Qy 390 --CAEGFAPIPH-----EPHRCQMFQACPADCDPNTQASCECPGY-I 432

Db 793 RQRCDGFLQDPGNCVDINECTSLSEPCRPGRGSCINTVGSYTCQRNPLI---CARGYHA 849

Qy 433 LDDGFICTDDECGNGFCSG---VCHNLPGTFECIC 466

Db 850 SDDGAKCVDNCECTGVHRCGEGQVCHNLPSYRCD 886

RESULT 8

A55567

fibrillin I - bovine

C:Species: Bos primigenius taurus (cattle)

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 02-Aug-2002

C:Accession: A55567

R:Tilstra, D.J.; Li, L.; Potter, K.A.; Womack, J.; Byers, P.H.

Genomics 23, 480-485, 1994

A:Title: Sequence of the coding region of the bovine fibrillin cDNA and localization to

A:Reference number: A55567; MUID:95137597; PMID:7835900

A:Accession: A55567

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-2871 <TIL>

A:Cross-references: GB:L28748; NID:g508427; PIDN:AAA74122.1; PID:g508428

C:Superfamily: fibrillin 1; EGF homology

F:1201-1236/Domain: EGF homology <EGF>

Query Match 12.6%; Score 367; DB 2; Length 2871;

Best Local Similarity 31.5%; Pred. No. 1.5e-14;

Matches 111; Conservative 26; Mismatches 109; Indels 106; Gaps 20;

Qy 244 DCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEFCVNP--- 300

Db 1200 ECSIMNGGCEFTCNSESYECSCQPGFALMPDQSC-----DIDE--CEDNPIC 1249

Qy 301 -----DQPSYSCMCTGYRLAQRHEDVDCCILEPSPC-PQRCVNTQGGFECHCYP 353

Db 1250 DGGQCTNIPGYRCLYDGFMASEDMDKTCVDVNECDLNPICLSGTCENTKGSFICHCDM 1309

Qy 354 NYDLVGE--CPEVPDPC--FRANCEYQCQPLNQT-SYLCVCAEFGA-----PIP 398

Db 1310 GYSKKGKTKGCTD-INCEIGAHCNDRHACTVNTAGSFKSCSPGWIQDGIKCTDLDECS 1368

Qy 399 HEPHRCQMFQACPADCDPNTQAS--CRCPGYILDDGFICTDIDE-EN----- 447

Db 1369 NGTHWCQH-----ADC-KNTMGSYRCLCKEY-TGDGFTCTDLDECSNLNCGNGQ 1419

Qy 448 -----GGF-----CS-----CVCHNLPGTFECICGPDALV 473

Db 1420 CLNAPGGYRCBDMGFVPSADGKACEDIDECSLFNICVFGTCHNLPLGRCEIGEYLD 1479

Qy 474 RHIG-----TDCDSK-VDGGDSGSGRPSPPTGSLTTPAVGLV 513

Db 1480 RSGNCNTDVNECDLPTTCISGNCVNTPGSYTCDCCPD-----FELNPTRVGCV 1527

RESULT 9

A54105

fibrillin-2 precursor - human

C:Species: Homo sapiens (man)

C>Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 02-Aug-2002

C:Accession: A54105; S17063; S31101

R:Zhang, H.; Apfelroth, S.D.; Hu, W.; Davis, E.C.; Sanguinetti, C.; Bonadio, J.; Mecham, J. Cell Biol. 124, 855-863, 1994

A:Title: Structure and expression of fibrillin-2, a novel microfibrillar component prefe

A:Reference number: A54105; MUID:94165150; PMID:8120105

A:Accession: A54105

Db 683 CKDNGPCKQVC---STVGGSAICSPPGYALMADGVSCEDINECVTDLHTCSRGHCNVT 739

Qy 344 QGGFECH---CYPNDYLDVGCVEPVPDPCFRANCEYQCQPLNQTSYLCV----- 389

Db 740 LGSFHCYKALTCEPGYALKDGC-EDVDEC--AMGTHTCQP-----GFLQCNKSGSYCOA 792

Qy 390 --CAEGFAPIPH-----EPHRCQMFQACPADCDPNTQASCECPGY-I 432

Db 793 RQRCDGFLQDPGNCVDINECTSLSEPCRPGRGSCINTVGSYTCQRNPLI---CARGYHA 849

Qy 433 LDDGFICTDDECGNGFCSG---VCHNLPGTFECIC 466

Db 850 SDDGAKCVDNCECTGVHRCGEGQVCHNLPSYRCD 886

RESULT 8

A55567

fibrillin I - bovine

C:Species: Bos primigenius taurus (cattle)

C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 02-Aug-2002

C:Accession: A55567

R:Tilstra, D.J.; Li, L.; Potter, K.A.; Womack, J.; Byers, P.H.

Genomics 23, 480-485, 1994

A:Title: Sequence of the coding region of the bovine fibrillin cDNA and localization to

A:Reference number: A55567; MUID:95137597; PMID:7835900

A:Accession: A55567

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-2871 <TIL>

A:Cross-references: GB:L28748; NID:g508427; PIDN:AAA74122.1; PID:g508428

C:Superfamily: fibrillin 1; EGF homology

F:1201-1236/Domain: EGF homology <EGF>

Query Match 12.6%; Score 367; DB 2; Length 2871;

Best Local Similarity 31.5%; Pred. No. 1.5e-14;

Matches 111; Conservative 26; Mismatches 109; Indels 106; Gaps 20;

Qy 244 DCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEFCVNP--- 300

Db 1200 ECSIMNGGCEFTCNSESYECSCQPGFALMPDQSC-----DIDE--CEDNPIC 1249

Qy 301 -----DQPSYSCMCTGYRLAQRHEDVDCCILEPSPC-PQRCVNTQGGFECHCYP 353

Db 1250 DGGQCTNIPGYRCLYDGFMASEDMDKTCVDVNECDLNPICLSGTCENTKGSFICHCDM 1309

Qy 354 NYDLVGE--CPEVPDPC--FRANCEYQCQPLNQT-SYLCVCAEFGA-----PIP 398

Db 1310 GYSKKGKTKGCTD-INCEIGAHCNDRHACTVNTAGSFKSCSPGWIQDGIKCTDLDECS 1368

Qy 399 HEPHRCQMFQACPADCDPNTQAS--CRCPGYILDDGFICTDIDE-EN----- 447

Db 1369 NGTHWCQH-----ADC-KNTMGSYRCLCKEY-TGDGFTCTDLDECSNLNCGNGQ 1419

Qy 448 -----GGF-----CS-----CVCHNLPGTFECICGPDALV 473

Db 1420 CLNAPGGYRCBDMGFVPSADGKACEDIDECSLFNICVFGTCHNLPLGRCEIGEYLD 1479

Qy 474 RHIG-----TDCDSK-VDGGDSGSGRPSPPTGSLTTPAVGLV 513

Db 1480 RSGNCNTDVNECDLPTTCISGNCVNTPGSYTCDCCPD-----FELNPTRVGCV 1527

RESULT 9

A54105

fibrillin-2 precursor - human

C:Species: Homo sapiens (man)

C>Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 02-Aug-2002

C:Accession: A54105; S17063; S31101

R:Zhang, H.; Apfelroth, S.D.; Hu, W.; Davis, E.C.; Sanguinetti, C.; Bonadio, J.; Mecham, J. Cell Biol. 124, 855-863, 1994

A:Title: Structure and expression of fibrillin-2, a novel microfibrillar component prefe

A:Reference number: A54105; MUID:94165150; PMID:8120105

A:Accession: A54105

A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual

A:Molecule type: mRNA

A:Residues: 1-2918 <ZHA>

A:Cross-references: GB:U03272

R:Lee, B.; Godfrey, M.; Vitale, E.; Hori, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras, Nature 352, 330-334, 1991

A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two diff

A:Reference number: S17062; MUID:91304567; PMID:1852206

A:Accession: S17063

A:Molecule type: mRNA

A:Residues: 752-1489, 1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P', 1928 <LEE>

A:Cross-references: EMBL:X62009

R:Milewicz, D.M.

submitted to the EMBL Data Library, December 1992

A:Reference number: S31101

A:Accession: S31101

A:Molecule type: mRNA

A:Residues: 752-1407, 'R', 1409-1489, 1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P',

A:Cross-references: EMBL:X62009

C:Genetics:

A:Gene: GDB:FBN2

A:Cross-references: GDB:L28122; OMIM:121050

A:Map position: 5q23-5q31

C:Superfamily: fibrillin 1; EGF homology

C:Keywords: extracellular protein

F:1-29/Domain: signal sequence #status predicted <SIG>

F:30-2918/Product: fibrillin-2 #status predicted <MAT>

F:1245-1280/Domain: EGF homology <EGF>

F:1970-2013/Domain: EGF homology <EGF>

Query Match 12.5%; Score 365; DB 2; Length 2918;

Best Local Similarity 33.1%; Pred. No. 2.1e-14;

Matches 101; Conservative 25; Mismatches 97; Indels 82; Gaps 17;

Qy 244 DCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEFCVNPDP-- 301

Db 1244 ECMIMNGGCDQCTNSSESYECSCSEGYALMPDGRSCA-----DIDE--CENNPDIC 1293

Qy 302 -----QPSYSCMCTGYRLAQRHEDVDCCILEPSPC-PQRCVNTQGGFECHCYP 353

Db 1294 DGGQCTNIPGYRCLYDGFMASEDMDKTCIDVNECDLNSNICMFEGECENTKGSFICHQCL 1353

Qy 354 NYDLVGE--CPEVPDPC--FRANCEYQCQPLNQT-SYLCVCAEFGAPIPH-----EPHR 403

Db 1354 GYSVKKGTKGCTD-VDECEIGAHCNMDHASCLNIPGSKSCREGW--INGIKCIDLDE 1410

Qy 404 COMFCNQACPADCDPNTQAS--CECPGYILDDGFICTDIDE-----CENG----- 448

Db 1411 CSNGTHQCSINAQC-VNTPGYSYRCACSEGF-TGDGFTCSDDVDECAENINLCENGQCLNVP 1468

Qy 449 -----GF-----CSGVCHNLPGTFECICGPDALVRHIG- 477

Db 1469 GAYRCECEMGTPASDSRSCQDIDECSPQNICVSGTCNNLPGMPHFCICDDGYELDRITGGN 1528

Qy 478 -TDCC 481

Db 1529 CTDID 1533

RESULT 10

A47221

fibrillin 1 precursor - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 02-Jun-1995 #sequence\_revision 25-Apr-1997 #text\_change 02-Aug-2002

C:Accession: A47221; I54355; S17064; I59574; S17062; S62111; A34198

R:Corson, G.M.; Chalberg, S.C.; Dietz, H.C.; Charbonneau, N.L.; Sakai, L.Y. Genomics 17, 476-484, 1993

A:Title: Fibrillin binds calcium and is coded by cDNAs that reveal a multidomain stru

A:Reference number: A47221; MUID:94010947; PMID:7691719

A:Accession: A47221

A:Molecule type: mRNA

A:Residues: 1-337, 'T', 339-1029 <COR>

A:Cross-references: GB:X63556

C:Species: Rattus norvegicus (Norway rat)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: T13954  
R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.  
Genomics 51, 27-34, 1998  
A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs  
A:Reference number: Z14126; MUID:98360089; PMID:9693030  
A:Accession: T13954  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1574 <NAK>  
A:Cross-references: EMBL:AB011532; NID:g3449293; PIDN:BAA32462.1; PID:g3449294  
A:Experimental source: strain Sprague-Dawley; brain  
C:Genetics:  
A:Gene: MBGF6

Query Match 14.2%; Score 414; DB 2; Length 1574;  
Best Local Similarity 34.1%; Pred. No. 1.3e-17;  
Matches 104; Conservative 28; Mismatches 101; Indels 72; Gaps 17;

QY 224 CTAPPG-AVGHWAREAPGAWDGSVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTA 282  
Db 147 CRPPGVQLQGD-GKTCQDVDECAHNGGQHRCVNTPGSYLCECKPGRFRLHTDGRCL- 204  
QY 283 SATQSC---NDLCEHFVCP-----NPDQPGSYSCM----- 309  
Db 205 -AISCTLNGGCGHQCQVLTQTQHRCCRPQYQLQEDGRRVRRSPCAEGNGGCMHICQ 263  
QY 310 -----CETGYRLAADQHRCDVDVDCILEPSPQRCVNTQGGEGECYCPNYDL-VD 359  
Db 264 ELRLAHCCHGPHGYQLAADRKTCDVDVDCALGLAQCAHGCCLNTQGSFKVCYCHAGYELGAD 323  
QY 360 G-EG-----VEPVDPCFRAN--CEVQCQPLNTSYLVCACAEGFAPIPHEPHRCQMF----- 408  
Db 324 GRQCYRTEMEIVNCEAGNGGCSHGCSH-TSTGPLCTCPRY-----ELDEDQKTCIDID 377  
QY 409 ---NOTACPADCDNTQA--SCEPEGYILD-DGFICTIDCEBNG-GFCSGVGNHNLPGT 461  
Db 378 DCANSPCCQAC-ANTPGYECSCFAGYRLNTDGGCEDVDGCASGGHGCHEHCSNLAGS 436  
QY 462 FECIC 466  
Db 437 FOCFC 441

RESULT 6  
A57278  
fibillin-2 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 23-Feb-1996 #sequence\_revision 23-Feb-1996 #text\_change 02-Aug-2002  
C:Accession: A57278  
R:Zhang, H.; Hu, W.; Ramirez, F.  
J. Cell Biol. 129, 1165-1176, 1995  
A:Title: Developmental expression of fibillin genes suggests heterogeneity of extracellular matrix  
A:Reference number: A57278; MUID:95263670; PMID:7744963  
A:Accession: A57278  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2907 <ZHA>  
A:Cross-references: GB:L39790; NID:g762830; PIDN:AAA74908.1; PID:g762831  
C:Superfamily: fibillin 1; EGF homology  
F:1239-1274/Domain: EGF homology <EGF1>  
F:2488-2523/Domain: EGF homology <EGF>

Query Match 12.8%; Score 373; DB 2; Length 2907;  
Best Local Similarity 32.9%; Pred. No. 6.8e-15;  
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;

QY 244 DCSVNGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFVCPNPD--- 301  
Db 1238 ECMIMNGGCDTQCTNSEGSEYCSCEGYALMPDGRSCA-----DIDE--CENNPDIC 1287  
QY 302 -----QPGSYSCMCTGYRLAADOHRCEDVDVDCILEPSPCP-QRCVNTQGGFECCHCYP 353

Db 1288 DGGQCTNIPGEYRCLCYDGFMAHMDKTCIDVNECDLNPNICMFGECENTKGSFICHQOL 1347  
QY 354 NYDLVDGE--CPEVDVPC--FRANCEYQCOPLN-QTSYLCVCAEGFA-----PIP 398  
Db 1348 GYSVKKCTTGCTD-VDECEIGAHCNDMHASCLNVPVGFKCSRGVNGVNGIKICIDLDECA 1406  
QY 399 HEPHRCQMFNCQACPADCPNTQAS--CECPEGYILLDDGFICTIDDE-----CENG- 448  
Db 1407 NGTHQCSI-----NAQC-VNTPGSRACACSEGF-TGDGFTCSVDVDECAENTNLCENGQ 1457  
QY 449 -----GF-----CS-----GVCHNLPGTFEFCICGPDLSALV 473  
Db 1458 CLNVPGAYRCBCEMGFTPASDSRSCQIDECFQNICVFGTCNNLPGMFHCICDDGYELD 1517  
QY 474 RHIG--TDCCD 481  
Db 1518 RTGGNCITDID 1527

RESULT 7  
A55184  
fibulin-2 precursor - human  
N:Alternate names: protein DKFZp586A1519.1  
C:Species: Homo sapiens (man)  
C:Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 02-Aug-2002  
C:Accession: A55184; T08744  
R:Zhang, R.Z.; Pan, T.C.; Zhang, Z.Y.; Mattei, M.G.; Timpl, R.; Chu, M.L.  
Genomics 22, 425-430, 1994  
A:Title: Fibulin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping of the gene  
A:Reference number: A55184; MUID:95104855; PMID:7806230  
A:Accession: A55184  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1184 <ZHA>  
A:Cross-references: GB:X82494; NID:g575232; PIDN:CAA57876.1; PID:g575233  
R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, March 1999  
A:Reference number: Z16471  
A:Accession: T08744  
A:Molecule type: mRNA  
A:Residues: 656-719; QDECLMGAHDCRRQFCVNTLGSFYCVNHTVLCADGYILNAHRKCDV', 720-853, 'T', 855  
A:Cross-references: EMBL:AL050095  
A:Experimental source: adult uterus; clone DKFZp586A1519  
C:Genetics:  
A:Gene: GDB:FBLN2  
A:Cross-references: GDB:293037; OMIM:135821  
A:Map position: 3p25-3p24  
A:Note: DKFZp586A1519.1  
C:Superfamily: fibulin-2; EGF homology  
C:Keywords: alternative splicing; extracellular matrix  
F:1-27/Domain: signal sequence status predicted <SIG>  
F:28-1184/Product: fibulin-2 protein #status predicted <MAT>  
F:905-941/Domain: EGF homology <EGF>

Query Match 12.6%; Score 367.5; DB 2; Length 1184;  
Best Local Similarity 30.0%; Pred. No. 6.7e-15;  
Matches 119; Conservative 29; Mismatches 124; Indels 125; Gaps 21;

QY 154 EBOQCEVKAD-GFLCEFHFPATC-----RPLAV-----EPGAAAASVITYTTPFAARG 201  
Db 531 EGQSCSNPLGYPCN-HVMLSCEGEPLTIVPEVRRPPEPAAAPRRVS-----EAEM 582  
QY 202 AFOQALPVGSSAAV---APLGQLMCTAPGAVQGHAREAPGAWDCSVNGGCEHACNA 258  
Db 583 AGREALSLGTEALPNSLPDQDQDECLLPFGL-----CQHLICIN 622  
QY 259 IGCAPRCQCPAGALQADGRSC-----TASAT-----QS 287  
Db 623 TVGSYHCACFPFGLSDGDTGRTCPREGHPPOPEAPQEPALKSEFSQVANSNTIPLPLPQNT 682  
QY 288 CND--LCEHFVCPNPDQPGSYSCMCTGYRLAADOHRCEDVDVDCILEPSPCP--QRCVNT 343

```

QY 349  CHCYPNYDLVDGECVEPDPFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMF 408
    |||  :|||:|||||  |||||:  ::  :|||:|:|  ||:|:|:|
Db 123  CHCDTGVELVDGECVDPDFCNFNCEYQCQPVGRSEHKICIAEGFAPVPGAPHKQMF 182
    |||  :|||:|||||  |||||:  ::  :|||:|:|  ||:|:|:|
QY 409  NQTACPADCPNTQASCEGPEGYLDGFGICTIDBCENGGFCGVCNHLPTGTFECIGP 468
    |||  :|||:|||||  |||||:  ::  :|||:|:|  ||:|:|:|
Db 183  NQTSACPADCPHYPTICRCPEGYIIDEGSTCTIDNECDT-NICPGQCNHLPTGYECIGP 241
    |||  :|||:|||||  |||||:  ::  :|||:|:|  ||:|:|:|
QY 469  DSALVRHIGWDCDSGKV-----DGDGSGGSEPPSPPTGSTLTP-PA-VGLVRSG 516
    ||||  ||  |||  :  |||  |||||  |||||:  ||  ||  :|||
Db 242  DSALSGQIGIDCDPTQVNEERGTPDYG--GSGEPPVSPPTGATRSPAPAGPLHSG 297
    ||||  ||  |||  :  |||  |||||  |||||:  ||  ||  :|||

RESULT 4
T27283
hypothetical protein Y64G10A.f - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T27283
R:Ainscough, R.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z20336
A:Accession: T27283
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1620 <WIL>
A:Cross-references: EMBL:AL110498; NID:el542303; PIDN:CAB54471.1; CBSP:Y64G10A
A:Experimental source: clone Y64G10A
C:Genetics:
A:Gene: CBSP:Y64G10A.f
A:Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 466/1; 548/1; 559/1; 601/1

Query Match 14.3%; Score 415; DB 2; Length 1620;
Best local similarity 27.0%; Pred. No. 1e-17;
Matches 128; Conservative 39; Mismatches 175; Indels 132; Gaps

```

[illegible]

Biochem. J. 295, 131-140, 1993  
A:Title: Identification of the predominant glycosaminoglycan-attachment site in soluble  
serine.  
A:Reference number: S38954; MUID:94029900; PMID:8216207  
A:Accession: S38954  
A:Molecule type: protein  
A:Residues: 475-491, 'X', 493-494 <GER>  
A:Note: the residue designated 'X' was determined to be a Ser with covalently bound chon  
R:Meininger, D.P.; Komives, E.A.  
submitted to the Brookhaven Protein Data Bank, September 1995  
A:Reference number: A67369; PDB:1ZQA  
A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue  
R:Tulinsky, A.; Mathews, I.I.  
submitted to the Brookhaven Protein Data Bank, August 1994  
A:Reference number: A52804; PDB:1HLT  
A:Contents: annotation; X-ray crystallography, 3.0 angstroms, residues 426-442  
R:Hrabal, R.; Komives, E.A.; Nl, F.  
submitted to the Brookhaven Protein Data Bank, November 1995  
A:Reference number: A65583; PDB:1FGD  
A:Contents: annotation; conformation by (1)H-NMR, residues 427-444  
R:Hrabal, R.; Komives, E.A.; Nl, F.  
Protein Sci. 5, 195-203, 1996  
A:Title: Structural resiliency of an EGF-like subdomain bound to its target protein, the  
A:Reference number: A58595; MUID:96276211; PMID:8745396  
A:Contents: annotation; conformation by (1)H-NMR  
C:Genetics:  
A:Gene: GDB:THBD  
A:Cross-references: GDB:119613; OMIM:188040  
A:Map position: 20p11.2-20p11.2  
A:Introns: #status absent  
C:Complex: homodimer, urinary form  
C:Function:  
A:Description: inhibits thrombin activation of fibrinogen; cofactor for thrombin activat  
A:Pathway: blood coagulation moderation  
A:Note: the membrane-bound form is located on the endothelium luminal surface of arterie  
A:Note: thrombin complexed with the membrane-bound form is subject to endocytosis  
C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology  
C:Keywords: anticoagulant; beta-hydroxyasparagine; beta-hydroxyaspartic acid; blood coag  
e protein  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-575/Product: thrombomodulin, membrane-bound form #status predicted <MAT>  
F:19-513/Domain: extracellular #status predicted <EXT>  
F:19-486/Product: thrombomodulin, urinary form #status experimental <MAU>  
F:24-167/Domain: C-type lectin homology <LCH>  
F:177-199/Region: PEST sequence  
F:201-233/Region: PEST sequence  
F:245-280/Domain: EGF homology <EG1>  
F:288-323/Domain: EGF homology <EG2>  
F:329-362/Domain: EGF homology <EG3>  
F:369-404/Domain: EGF homology <EG4>  
F:408-439/Domain: EGF homology <EG5>  
F:445-480/Domain: EGF homology <EG6>  
F:485-513/Region: PEST sequence  
F:517-539/Domain: transmembrane #status predicted <TMN>  
F:540-575/Domain: intracellular #status predicted <INT>  
F:47,115,116,382,409/Binding site: carbonylate (Asn) (covalent) #status predicted  
F:174,225,411,504/Binding site: carbonylate (Thr) (covalent) #status predicted  
F:245-256,252-267,280-288,292-308,310-323,329-340,336-349,351-362,369-378,374-38  
F:334,498/Binding site: carbonylate (Ser) (covalent) #status predicted  
F:342/Modified site: erythro-beta-hydroxyasparagine (Asn) #status experimental  
F:490,492/Binding site: Chondroitin sulfate (Ser) (covalent) (partial) #status experimen

Query Match 100.0%; Score 2916; DB 1; Length 575;  
Best Local Similarity 100.0%; Pred. No. 5.1e-169;  
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVVLVGLALAGLGFAPAPPGGSGQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
DB 1 MGVVLVGLALAGLGFAPAPPGGSGQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVISLLNGDGGVRRRLWGLQLPPGCGPKRLGPGFQWVTGDNNTSYS 120  
DB 61 TVRSSVAADVISLLNGDGGVRRRLWGLQLPPGCGPKRLGPGFQWVTGDNNTSYS 120

QY 121 RWARLDLNGAPLCGPPLCVAVSAEATVPSEPITWEEOQCEVKADGFLCEFHFPATCRPLAV 180  
DB 121 RWARLDLNGAPLCGPPLCVAVSAEATVPSEPITWEEOQCEVKADGFLCEFHFPATCRPLAV 180  
QY 181 EPGAAAAVSIITGTPFAARGADFOALPYGSSAAVAPLGLQMLCTAPPAGVAGVGHAREAP 240  
DB 181 EPGAAAAVSIITGTPFAARGADFOALPYGSSAAVAPLGLQMLCTAPPAGVAGVGHAREAP 240  
QY 241 GAWDCSVENGSGCCEHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
DB 241 GAWDCSVENGSGCCEHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DQPGSYSCMCETGYRLAADOHRCEDVDDCILLESPPCQRCVNTQGGFECHECHYNYDLVDG 360  
DB 301 DQPGSYSCMCETGYRLAADOHRCEDVDDCILLESPPCQRCVNTQGGFECHECHYNYDLVDG 360  
QY 361 ECVPEVDPGCFRANCEYOCQPLNOTSYLVCACGFAFIPHEPHRCOMFCNOTACPADCDPN 420  
DB 361 ECVPEVDPGCFRANCEYOCQPLNOTSYLVCACGFAFIPHEPHRCOMFCNOTACPADCDPN 420  
QY 421 TQASCEPGEYILDDGFICTDIDECENGSGFCVCHNLPGTFECICGPDALSALVHRHIGTDC 480  
DB 421 TQASCEPGEYILDDGFICTDIDECENGSGFCVCHNLPGTFECICGPDALSALVHRHIGTDC 480  
QY 481 DSGKVDGDSGSGEPPSPPTPGSTLTTPPAVGLVHSG 516  
DB 481 DSGKVDGDSGSGEPPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 2  
A60501  
Thrombomodulin precursor - mouse  
N:Alternate names: fetomodulin  
C:Species: Mus musculus (house mouse)  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 16-Jul-1999  
C:Accession: S08488; A32001; A60501  
R:Dittman, W.A.; Majerus, P.W.  
Nucleic Acids Res. 17, 802, 1989  
A:Title: Sequence of a cDNA for mouse thrombomodulin and comparison of the predicted  
A:Reference number: S08488; MUID:89128454; PMID:2536925  
A:Accession: S08488  
A:Molecule type: mRNA  
A:Residues: 1-577 <DIT>  
A:Cross-references: EMBL:X14432; NID:g54781; PIDN:CAA32597.1; PID:g54782  
R:Dittman, W.A.; Kumada, T.; Sadler, J.E.; Majerus, P.W.  
J. Biol. Chem. 263, 15815-15822, 1988  
A:Title: The structure and function of mouse thrombomodulin. Phorbol myristate acetat  
A:Reference number: A32001; MUID:89008498; PMID:2844823  
A:Accession: A32001  
A:Molecule type: mRNA  
A:Residues: 97-577 <D12>  
A:Cross-references: GB:J04060  
R:Imada, S.; Yamaguchi, H.; Nagumo, M.; Katayanagi, S.; Iwasaki, H.; Imada, M.  
Dev. Biol. 140, 113-122, 1990  
A:Title: Identification of fetomodulin, a surface marker protein of fetal development  
A:Reference number: A60501; MUID:90292331; PMID:2162790  
A:Accession: A60501  
A:Molecule type: protein  
A:Residues: 19-22;330-343;479-489;545-555;562-575 <IMA>  
C:Comment: Thrombomodulin binds to and internalizes with thrombin. It is also a cofac  
C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology  
C:Keywords: anticoagulant; membrane protein; phosphoprotein; receptor  
F:24-165/Domain: C-type lectin homology <LCH>  
F:287-322/Domain: EGF homology <EG1>  
F:328-361/Domain: EGF homology <EG2>  
F:368-403/Domain: EGF homology <EG3>  
F:407-438/Domain: EGF homology <EG4>  
F:444-479/Domain: EGF homology <EG5>

Query Match 65.1%; Score 1899; DB 2; Length 577;  
Best Local Similarity 65.5%; Pred. No. 1.3e-107;

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:22:58 ; Search time 16.5 seconds

(without alignments)  
3006.386 Million cell updates/sec

Title: US-09-509-994-1

Perfect score: 2916

Sequence: 1 MLGVLVGLALAGLGPAP.....PSPTPGSLTPPAVLVHSG 516

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 9613422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_73.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2916	100.0	575	1 THHUB	thrombomodulin pre
2	1899	65.1	577	2 A60501	thrombomodulin pre
3	1092.5	37.5	356	2 A25918	thrombomodulin - b
4	416	14.3	1620	2 T27283	hypothetical prote
5	414	14.2	1574	2 T13954	MEGF6 protein - ra
6	373	12.8	2907	2 A52728	fibrillin-2 precu
7	367.5	12.6	1184	2 A55184	fibulin-2 precuso
8	367	12.6	2871	2 A55567	fibrillin 1 - bovi
9	365	12.5	2918	2 A54105	fibrillin-2 precu
10	363	12.4	3002	2 A47221	fibrillin-1 precu
11	361	12.4	2871	2 A55624	fibrillin-1 precu
12	359	12.3	741	2 T46488	hypothetical prote
13	357	12.2	1221	2 T49457	fibulin-2 precuso
14	331	11.4	589	2 T43210	fibulin-1D precurs
15	329.5	11.3	1964	2 T09059	notch4 - mouse
16	327.5	11.2	1712	2 A38261	masking protein pr
17	324.5	11.1	689	2 T42760	fibulin, splice fo
18	324.5	11.1	712	2 T42990	fibulin 1, splice
19	320.5	11.0	2321	2 T78549	notch3 protein - h
20	320	11.0	1394	2 A35626	transforming grow
21	314.5	10.8	3507	2 T34513	hypothetical prote
22	313	10.7	601	2 B36346	fibulin 1 precuso
23	313	10.7	683	2 C36346	fibulin 1 precuso
24	313	10.7	798	2 T22793	hypothetical prote
25	312.5	10.7	1820	2 A55494	latent transformin
26	308	10.6	685	2 T78040	fibulin, splice fo
27	308	10.6	705	2 S34968	fibulin, splice fo
28	303.5	10.4	1251	2 A57293	latent transformin
29	296	10.2	2531	2 A46019	Notch-1 protein -

#### ALIGNMENTS

##### RESULT 1

##### THHUB

thrombomodulin precursor [validated] - human

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1988 #sequence\_revision 12-May-1995 #text\_Change 15-Sep-2000

C:Accession: A41442; A28307; A29680; A27073; JX0264; S38954

R:Shirai, T.; Shiojiri, S.; Ito, H.; Yamamoto, S.; Kusumoto, H.; Deyashiki, Y.; Maruy

J. Biochem. 103, 281-285, 1988

A:Title: Gene structure of human thrombomodulin, a cofactor for thrombin-catalyzed ac

A:Reference number: A41442; MUID:88227901; PMID:2836377

A:Accession: A41442

A:Molecule type: DNA

A:Residues: 1-575 <SHI>

A:Cross-references: DDBJ:D00210; NID:g220126; PIDN:BAA00149.1; PID:g220127

R:Jackman, R.W.; Beeler, D.L.; Fritze, L.; Soff, G.; Rosenberg, R.D.

Proc. Natl. Acad. Sci. U.S.A. 84, 6425-6429, 1987

A:Title: Human thrombomodulin gene is intron depleted: nucleic acid sequences of the

A:Reference number: A28307; MUID:87317665; PMID:2819876

A:Accession: A28307

A:Molecule type: DNA; mRNA

A:Residues: 1-472, 'A', 474-575 <JAC>

A:Cross-references: GB:J02973; NID:g339658; PIDN:AAA61175.1; PID:g339659

R:Suzuki, K.; Kusumoto, H.; Deyashiki, Y.; Nishioka, J.; Maruyama, I.; Zushi, M.; Kaw

EMBO J. 6, 1891-1897, 1987

A:Title: Structure and expression of human thrombomodulin, a thrombin receptor on end

A:Reference number: A29680; MUID:88004395; PMID:2820710

A:Accession: A29680

A:Molecule type: mRNA

A:Residues: 1-575 <SUZ>

A:Cross-references: GB:X05495; NID:g37123; PIDN:CAA29045.1; PID:g736251

A:Experimental source: lung endothelium

A:Note: part of this sequence, including the amino end of the mature protein, were de

R:Wen, D.; Dittman, W.A.; Ye, R.D.; Deaven, L.L.; Majerus, P.W.; Sadler, J.E.

Biochemistry 26, 4350-4357, 1987

A:Title: Human thrombomodulin: complete cDNA sequence and chromosome localization of

A:Reference number: A27073; MUID:88024950; PMID:2822087

A:Accession: A27073

A:Molecule type: mRNA

A:Residues: 1-472, 'A', 474-575 <WEN>

A:Cross-references: GB:M16552; NID:g339656; PIDN:AAB59508.1; PID:g339657

A:Experimental source: placenta

A:Note: parts of this sequence were determined by protein sequencing

R:Yamamoto, S.; Mizoguchi, T.; Tamaki, T.; Ohkuchi, M.; Kimura, S.; Aoki, N.

J. Biochem. 113, 433-440, 1993

A:Title: Urinary thrombomodulin, its isolation and characterization.

A:Reference number: JX0264; MUID:93293792; PMID:8390445

A:Accession: JX0264

A:Molecule type: protein; mRNA

A:Residues: 19-472, 'A', 474-486 <YAM>

A:Experimental source: urine

A:Note: the urinary form appears to be identical with that circulating in plasma

R:Gerlitz, B.; Hassell, T.; Vlahos, C.J.; Parkinson, J.F.; Bang, N.U.; Grinnell, B.W.



```

XX Recombinant protein having thrombomodulin activity -
PT used in the therapeutic control of coagulation and the treatment
PT and prevent of thrombotic episodes
XX
PS Disclosure; ; p; English.
XX
CC Protein is encoded by plasmid p2.1 derived from human cell
CC line A549 known to express about 10000 molecules of thrombomodulin
CC per cell. Plasmid p2.1 showed a strong hybridisation signal with a
CC 60-mer bovine thrombomodulin probe. Thrombomodulins are used to
CC potentiate a patient's anticoagulant capacity.
CC See also AAN82027 and AAN82037.
XX
SQ Sequence 575 AA;

Query Match          99.7%; Score 2906; DB 9; Length 575;
Best Local Similarity 99.6%; Pred. No. 1.7e-161;
Matches 514; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MGVLVLGALALAGLPAPAPQPGSQVHCDFALYPGPATFLNASQICDGLRGLHM 60
DB 1 MGVLVLGALALAGLPAPAPQPGSQVHCDFALYPGPATFLNASQICDGLRGLHM 60

QY 61 TVRSSVAADVLSLLNGDGGVGRRLWIGLQLPGCCGDPKRLGPIRGFOWVTGDNNTSYS 120
DB 61 TVRSSVAADVLSLLNGDGGVGRRLWIGLQLPGCCGDPKRLGPIRGFOWVTGDNNTSYS 120

QY 121 RWARDLNGALPCGLCVAVSAEAATVSEPIWEEQQCEVKADGFLCEFFHPATCRPLAV 180
DB 121 RWARDLNGALPCGLCVAVSAEAATVSEPIWEEQQCEVKADGFLCEFFHPATCRPLAV 180

QY 181 EPGAAAASVITYGTPFAARGADQALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPFAARGADQALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240

QY 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPNP 300
DB 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVPNP 300

QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDG 360
DB 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTQGGFECHCYPNYDLVDG 360

QY 361 ECVPEVDFCFRANCYQCPQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420
DB 361 ECVPEVDFCFRANCYQCPQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420

QY 421 TQASCEPEGYILDGFICTDIDECENGFCGVCYCHNLPGTFECICGPDLSALVRHIGTDC 480
DB 421 TQASCEPEGYILDGFICTDIDECENGFCGVCYCHNLPGTFECICGPDLSALVRHIGTDC 480

QY 481 DSGKVDGSDSGSGEPPTPGSTLTTPAVGLVHSG 516
DB 481 DSGKVDGSDSGSGEPPTPGSTLTTPAVGLVHSG 516
```

Search completed: December 16, 2002, 17:24:24  
Job time : 37.5 secs



Db 301 DQGSYSCHMETGYRLAADOHRCEVDVDCILPSPQRCVNTQGGFECHECHYPNYDLVDG 360  
Qy 361 ECVEPVDFPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420  
Db 361 ECVEPVDFPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420  
Qy 421 TQASCEPCEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDC 480  
Db 421 TQASCEPCEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDC 480  
Qy 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPAVGLVHSG 516  
Db 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPAVGLVHSG 516

RESULT 14  
AAR10617  
ID AAR10617 standard; Protein; 515 AA.  
XX  
AC AAR10617;  
DT 17-APR-1991 (first entry)  
XX  
DE Soluble thrombomodulin deriv.  
KW Thrombosis; anticoagulant; Protein C.  
XX  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Peptide 1..16  
FT /label= signal peptide  
FT /note= "for low mol. wt deriv.; optionally present"  
FT Peptide 1..18  
FT /label= signal peptide  
FT /note= "for high mol wt. deriv; optionally present"  
XX  
EP412841-A.  
XX  
PD 13-FEB-1991.  
XX  
PF 10-AUG-1990; 90EP-0308826.  
XX  
PR 05-FEB-1990; 90US-0474870.  
PR 11-AUG-1989; 89US-0393617.  
XX  
PA (ELIL ) ELI LILLY & CO.  
XX  
PI Bang NU, Grinnell BW, Hoskins JA, Moore RE, Parkinson JF;  
XX  
XX WPI; 1991-045960/07.  
XX N-PSDB; AAQ10435.  
PT New soluble thrombomodulin derivatives - produced by recombinant  
PT DNA techniques for use as anticoagulants and in treatment of  
PT thrombosis.  
XX  
PS Claim 2; Page 45; 81pp; English.

CC The sequence is that of a sol. thrombomodulin deriv. comprising the  
CC signal peptide (optional), N-terminal, epidermal growth factor  
CC homology region and the Ser/Thr-rich region, but lacking the trans-  
CC membrane and cytoplasmic domains. The sequence was deduced DNA  
CC obtd. from clone GHTW3A isolated from a human chromosome 20 library,  
CC subcloned into pUC19 to give pGHTW3A. A Ppu MI fragment of this  
CC vector was ligated with a linker to construct pUC18TM, which was  
CC treated with BsmI to delete about 500 bp from the 3' end of the  
CC gene. ( The deletion occurs at the Ser/Thr-rich/transmembrane domain  
CC junction, effectively cleaving the transmembrane and cytoplasmic  
CC domains.) The plasmid was recircularised to give pUC18TM for use  
CC in prodn of the recombinant sol. deriv. The protein reacts with  
CC thrombin to activate the Protein C anticoagulant pathway, inhibit  
CC thrombin, and can be used to treat or prevent thrombotic disorders.

XX SQ Sequence 515 AA;  
Query Match 99.7%; Score 2906; DB 12; Length 515;  
Best Local Similarity 99.8%; Pred. No. 1.5e-161;  
Matches 514; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MGVLVLCALALAGLGFAPAPAPQPGGSCQVEHDCFALYPGPATFLNASQICDGLRHLM 60  
Db 1 MGVLVLCALALAGLGFAPAPAPQPGGSCQVEHDCFALYPGPATFLNASQICDGLRHLM 60  
Qy 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQLPCCGDPKRLGRLGFGQWVTGDNNTSYS 120  
Db 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQLPCCGDPKRLGRLGFGQWVTGDNNTSYS 120  
Qy 121 RWARDLNGAPLCGPLCVAVSAAEATVPSEPIWEQQQCEVKADGFLCEHFHFCRPLAV 180  
Db 121 RWARDLNGAPLCGPLCVAVSAAEATVPSEPIWEQQQCEVKADGFLCEHFHFCRPLAV 180  
Qy 181 EPGAAAAVSTYGTTPFAARGADFOALPVGSSAAVAPLGQLMCTAPPAGVQGHWAREAP 240  
Db 181 EPGAAAAVSTYGTTPFAARGADFOALPVGSSAAVAPLGQLMCTAPPAGVQGHWAREAP 240  
Qy 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
Qy 301 DQGSYSCHMETGYRLAADOHRCEVDVDCILPSPQRCVNTQGGFECHECHYPNYDLVDG 360  
Db 301 DQGSYSCHMETGYRLAADOHRCEVDVDCILPSPQRCVNTQGGFECHECHYPNYDLVDG 360  
Qy 361 ECVEPVDFPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420  
Db 361 ECVEPVDFPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420  
Qy 421 TQASCEPCEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDC 480  
Db 421 TQASCEPCEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDC 480  
Qy 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPAVGLVHSG 515  
Db 481 DSGKVDGSDSGSGEPSPPTPGSTLTPPAVGLVHSG 515  
RESULT 15  
AAP82070  
ID AAP82070 standard; protein; 575 AA.  
XX  
AC AAP82070;  
XX  
DT 19-OCT-1990 (first entry)  
XX  
DE Human thrombomodulin encoded by plasmid p2.1.  
XX  
KW thrombomodulin activity; Protein C; anticoagulant;  
KW epidermal Growth Factor (EGF) domains.  
XX  
OS synthetic.  
XX  
XX WO8809811-A.  
XX  
PD 15-DEC-1988.  
XX  
PF 09-JUN-1988; 88WO-DK00089.  
XX  
PR 12-JUN-1987; 87DK-0002990.  
XX  
PA (NOVO ) NOVO INDUSTRI A/S.  
XX  
PI Nexo BA, Esper B;  
XX  
XX WPI; 1988-368626/51.  
DR  
DR N-PSDB; AAP82026.

XX 28-JUL-2000 (first entry)  
 XX Human thrombomodulin TMD protein #2.  
 DE Human; thrombomodulin; vasculitis; protein C; thrombin.  
 XX Homo sapiens.  
 XX JP2000053582-A.  
 XX 22-FEB-2000.  
 XX 06-AUG-1998; 98JP-0222688.  
 XX 06-AUG-1998; 98JP-0222688.  
 XX (ASAH ) ASahi Kasei Kogyo KK.  
 XX WPI; 2000-353249/31.  
 XX N-PSDB; AAA10028.  
 XX Treating agent for vasculitis contains peptide which promotes  
 PT activation of protein C by thrombin -  
 XX Claim 4; Page 13-14; 18pp; Japanese.  
 XX This sequence represents a human thrombomodulin protein. The invention  
 CC relates to a treating agent for vasculitis containing a peptide which  
 CC promotes activation of protein C by thrombin.  
 XX SQ Sequence 516 AA;

Query Match 99.9%; Score 2912; DB 21; Length 516;  
 Best Local Similarity 99.8%; Pred. No. 6.8e-162;  
 Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MGVLVLGALAGLGPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHM 60  
 DB 1 MGVLVLGALAGLGPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHM 60  
 QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPCCGDPKRLGRLGFQWVTGDNNTSYS 120  
 DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPCCGDPKRLGRLGFQWVTGDNNTSYS 120  
 QY 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAV 180  
 DB 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAV 180  
 QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
 DB 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
 QY 241 GAWDCSVENGCGEACNATPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVPNP 300  
 DB 241 GAWDCSVENGCGEACNATPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVPNP 300  
 QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNYDLVDG 360  
 DB 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNYDLVDG 360  
 QY 361 ECVEPVDFCFRANCYQCPQLNQTSLYLCVCAEGFAPIPHEPHRCQFNCQACPADCDPN 420  
 DB 361 ECVEPVDFCFRANCYQCPQLNQTSLYLCVCAEGFAPIPHEPHRCQFNCQACPADCDPN 420  
 QY 421 TQASCECPGYILDDGFICTDIDECENGCGFCGVCNLPPTGFTFCICGPDALVRHIGTDC 480  
 DB 421 TQASCECPGYILDDGFICTDIDECENGCGFCGVCNLPPTGFTFCICGPDALVRHIGTDC 480  
 QY 481 DSGKVDGDSGSGEPPPTGFTTTPPAVGLVHSG 516  
 DB 481 DSGKVDGDSGSGEPPPTGFTTTPPAVGLVHSG 516

RESULT 13  
 AAR22017  
 ID AAR22017 standard; Protein; 516 AA.  
 XX  
 XX AAR22017;  
 XX 03-JUL-1992 (first entry)  
 XX Human thrombomodulin (1-516) with Asp367 substituted by Ala.  
 DE Mutant; thrombin binding site; blood clotting; Tmml mutator.  
 XX Homo sapiens.  
 XX EP474273-A.  
 XX 11-MAR-1992.  
 XX 05-AUG-1991; 91EP-0202009.  
 XX 03-AUG-1990; 90JP-0204978.  
 XX (ASAH ) ASahi Kasei Kogyo.  
 XX Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;  
 XX WPI; 1992-081820/11.  
 XX New polypeptide inhibits blood coagulation and platelet  
 PT aggregation - promotes thrombin catalysed protein C activation  
 PT for treating myocardial infarction, thrombosis, embolism, etc.  
 XX Example 1; Page 18; 112pp; English.  
 XX Plasmid M13TMD1 (see AAR22016) encoding the first 516 N-terminal  
 CC amino acids of human thrombomodulin was used as a template for  
 CC site-directed mutagenesis using the "Tmml mutator" to produce  
 CC plasmid M13TMD1. In the mutant thrombomodulin encoded by the  
 CC plasmid, the wild-type Asp residue at position 367 is substituted  
 CC by an Ala residue. This mutant, truncated thrombomodulin has low  
 CC activity in the activation of protein C, c.f. high activity for the  
 CC similarly truncated thrombomodulin with Asp 367.  
 CC See AAR22013-R22022 and AAQ25073.

Query Match 99.7%; Score 2908; DB 13; Length 516;  
 Best Local Similarity 99.8%; Pred. No. 1.2e-161;  
 Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MGVLVLGALAGLGPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHM 60  
 DB 1 MGVLVLGALAGLGPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHM 60  
 QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPCCGDPKRLGRLGFQWVTGDNNTSYS 120  
 DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPCCGDPKRLGRLGFQWVTGDNNTSYS 120  
 QY 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAV 180  
 DB 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCFEHPATCRPLAV 180  
 QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
 DB 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
 QY 241 GAWDCSVENGCGEACNATPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVPNP 300  
 DB 241 GAWDCSVENGCGEACNATPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVPNP 300  
 QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNYDLVDG 360  
 DB 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFCHCYPNYDLVDG 360

Matches 515; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVVLGALALAGLGPAPAEPPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60  
DB 1 MLGVVLGALALAGLGPAPAEPPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPPCGDKPKRLGLRGLFQWVGTGNNNTSYS 120  
DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPPCGDKPKRLGLRGLFQWVGTGNNNTSYS 120  
QY 121 RWARLDLNGAPLCGPLCVAVSAEATVPSEPIWEEQCEVKADGFLCEFFHPPATCRPLAV 180  
DB 121 RWARLDLNGAPLCGPLCVAVSAEATVPSEPIWEEQCEVKADGFLCEFFHPPATCRPLAV 180  
QY 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
DB 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
DB 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPCPQRCVNTQGGFECCHYPNYDLVDG 360  
DB 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPCPQRCVNTQGGFECCHYPNYDLVDG 360  
QY 361 ECVEPVPDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420  
DB 361 ECVEPVPDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420  
QY 421 TOASCCEPEGIYLDGFICTDIDECENGCGFCGVCVCHNLPGTFECICGPDSSALVRHIGTDC 480  
DB 421 TOASCCEPEGIYLDGFICTDIDECENGCGFCGVCVCHNLPGTFECICGPDSSALVRHIGTDC 480  
QY 481 DSGKVDGSDSGSGEPPTGSLTTPPAVGLVHSG 516  
DB 481 DSGKVDGSDSGSGEPPTGSLTTPPAVGLVHSG 516

RESULT 11  
AAY09348  
ID AAY09348 standard; Protein: 516 AA.  
XX  
AC AAY09348;  
XX  
DT 08-JUL-1999 (first entry)  
XX  
DE Human thrombomodulin SEQ ID NO:2.  
XX  
KW Human; thrombomodulin; aqueous parenteral solution; storage;  
KW distribution; acute coronary syndrome; thrombosis; embolism;  
KW diabetes.  
XX  
OS Homo sapiens.  
XX  
PN WO9918994-A1.  
XX  
PD 22-APR-1999.  
XX  
PF 13-OCT-1998; 98WO-JP04609.  
XX  
PR 11-NOV-1997; 97JP-0308523.  
PR 15-OCT-1997; 97JP-0281659.  
XX  
PA (ASAH ) ASAH KASEI KOGYO KK.  
XX  
PI Murata T, Shimizu H, Tsuruta K, Yokozawa A, Yui M;  
XX  
DR WPI; 1999-277444/23.  
DR N-PSDB; AAX53680.  
XX  
PT Stable aqueous parenteral thrombomodulin solution - comprising  
PT buffer and surfactant, useful for treating acute coronary syndrome,

PT thrombosis, embolism, and diabetes  
XX  
PS Claim 7; Page 90-92; 97pp; Japanese.  
XX  
CC The present invention describes a method for maintaining the quality of  
CC an aqueous parenteral solution of thrombomodulin comprising buffer and  
CC surfactant aseptically filled in a case or syringe. Maintaining the  
CC quality of an aqueous, parenteral thrombomodulin solution is  
CC characterised in that the solution: (a) comprises soluble thrombomodulin,  
CC a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically  
CC filled into: (i) a case or (ii) a syringe without any empty space; and  
CC (c) is kept in liquid form in storage and distribution and not frozen or  
CC freeze dried. Thrombomodulin is used to treat acute coronary syndrome  
CC (such as myocardial infarction, unstable angina and coronary artery  
CC blockage), thrombosis (e.g. cerebral, vascular and peripheral blood  
CC vessel thrombosis), embolism, peripheral blood vessel disorders (e.g.  
CC Raynauds disease), arteriosclerosis, vasculitis (e.g. due to systemic  
CC lupus erythematosus or Barretts syndrome), multiple organ failure,  
CC disseminated intravascular coagulation, transient ischaemia, diabetes,  
CC liver veno-occlusive diseases and deep vein thrombosis. The composition  
CC is stable for a long period of time and can be stored and distributed in  
CC ready to use form avoiding the problems of dissolution and accuracy when  
CC preparing on demand. The present sequence represents human  
CC thrombomodulin.  
XX  
SQ Sequence 516 AA;

Query Match 99.9%; Score 2912; DB 20; Length 516;  
Best Local Similarity 99.8%; Pred. No. 6.8e-162;  
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGVVLGALALAGLGPAPAEPPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60  
DB 1 MLGVVLGALALAGLGPAPAEPPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPPCGDKPKRLGLRGLFQWVGTGNNNTSYS 120  
DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPPCGDKPKRLGLRGLFQWVGTGNNNTSYS 120  
QY 121 RWARLDLNGAPLCGPLCVAVSAEATVPSEPIWEEQCEVKADGFLCEFFHPPATCRPLAV 180  
DB 121 RWARLDLNGAPLCGPLCVAVSAEATVPSEPIWEEQCEVKADGFLCEFFHPPATCRPLAV 180  
QY 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
DB 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240  
QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
DB 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPCPQRCVNTQGGFECCHYPNYDLVDG 360  
DB 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPCPQRCVNTQGGFECCHYPNYDLVDG 360  
QY 361 ECVEPVPDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420  
DB 361 ECVEPVPDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420  
QY 421 TOASCCEPEGIYLDGFICTDIDECENGCGFCGVCVCHNLPGTFECICGPDSSALVRHIGTDC 480  
DB 421 TOASCCEPEGIYLDGFICTDIDECENGCGFCGVCVCHNLPGTFECICGPDSSALVRHIGTDC 480  
QY 481 DSGKVDGSDSGSGEPPTGSLTTPPAVGLVHSG 516  
DB 481 DSGKVDGSDSGSGEPPTGSLTTPPAVGLVHSG 516

RESULT 12  
AAY83935  
ID AAY83935 standard; Protein: 516 AA.  
XX  
AC AAY83935;

FT Domain 19..244  
 FT /label= N-terminal\_domain  
 FT 245..480  
 FT /note= "6 EGF-like domains"  
 FT Modified-site 481..515  
 FT /note= "O-linked glycosylation site"  
 FT 516..539  
 FT /label= Transmembrane\_domain  
 FT 540..575  
 FT /label= Cytoplasmic\_domain  
 PN W0200198352-A2.  
 XX 27-DEC-2001.  
 XX 20-JUN-2001; 2001WO-US19590.  
 XX 21-JUN-2000; 2000US-213678P.  
 PR 12-JUN-2001; 2001US-0880484.  
 XX (SCHD ) SCHERING AG.  
 XX Light D, Morser MJ, Nagashima M;  
 PI N-PSDB; AAD28182.  
 DR WPI; 2002-154622/20.  
 DR N-PSDB; AAD28182.  
 XX Thrombomodulin analog with greater ability to potentiate  
 PT thrombin-mediated activation of protein C and lesser ability to  
 PT potentiate thrombin-mediated activation of thrombin-activatable  
 PT fibrinolysis inhibitor for treating thrombotic disease -  
 XX Claim 1; Fig 4; 53pp; English.  
 PS The invention relates to the design, production and use of thrombomodulin  
 XX (TM) analogues that have the ability to enhance the thrombin-mediated  
 CC activation of protein C which have a significantly reduced ability to  
 CC promote activation of thrombin-activatable fibrinolysis inhibitor (TAFI).  
 CC TM analogue is useful for treating thrombotic disease in human, by  
 CC administering it or a polynucleotide encoding it to the patient and  
 CC expressing the analogue in vivo. Composition comprising TM analogue is  
 CC useful for treating a thrombotic disease or condition in mammals. TM  
 CC analogue is useful in antithrombotic therapy. TM analogue is also useful  
 CC for treating diseases in which thrombus formation plays a significant  
 CC etiological role including myocardial infarction, disseminated  
 CC intravascular coagulation (DIC), deep vein thrombosis, septic shock,  
 CC pulmonary embolism, acute respiratory distress syndrome, unstable angina  
 CC and other arterial or venous occlusive conditions. TM analogue is also  
 CC useful for treating various circulatory disorders including coronary or  
 CC pulmonary embolism, stroke and systemic coagulation disorders including  
 CC DIC which is often associated with septicemia, certain cancers and  
 CC toxemia of pregnancy. The present sequence is human full-length native  
 CC thrombomodulin.  
 XX Sequence 575 AA;  
 SQ Query Match 100.0%; Score 2916; DB 23; Length 575;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-162;  
 Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MGLVVLGALALAGLGFPAPEPQGGSCVHEHDCFALYPGPATFNASQICDGLRGHLM 60  
 DB 1 MGLVVLGALALAGLGFPAPEPQGGSCVHEHDCFALYPGPATFNASQICDGLRGHLM 60  
 QY 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQLPPCGDPRKLGRLGFWQVNTGNTSYS 120  
 DB 61 TVRSSVAADVISILLNGDGGVRRRLWIGLQLPPCGDPRKLGRLGFWQVNTGNTSYS 120  
 QY 121 RWARLDNGLAPLGGPLCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEHFPATCRPLAV 180  
 DB 121 RWARLDNGLAPLGGPLCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEHFPATCRPLAV 180  
 QY 181 EPGAAAAVSIYGTPTFAARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240

Db 181 EPGAAAAVSIYGTPTFAARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240  
 QY 241 GAWDCSVNGGCEHACNAIPGAPRCQCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
 Db 241 GAWDCSVNGGCEHACNAIPGAPRCQCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
 QY 301 DOPGSYSCMCEYGLAADOHRCEVDVDCILEPSPCQRCVNTQGGFEGCHCYPNYDLVDG 360  
 Db 301 DOPGSYSCMCEYGLAADOHRCEVDVDCILEPSPCQRCVNTQGGFEGCHCYPNYDLVDG 360  
 QY 361 ECVEPVDFCFRANCEYQCQPLNOTSYLCVCAEGFAPIDPEHRCOMFCNOTACPADCDPN 420  
 Db 361 ECVEPVDFCFRANCEYQCQPLNOTSYLCVCAEGFAPIDPEHRCOMFCNOTACPADCDPN 420  
 QY 421 TOASCECEGYLLDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALVRHIGTDC 480  
 Db 421 TOASCECEGYLLDDGFICTDIDECENGFCGVCNHLPGTFECICGPDALVRHIGTDC 480  
 QY 481 DSGKVDGDSGSGEBPPSPPTPGSTLTTPPAVGLVHSG 516  
 Db 481 DSGKVDGDSGSGEBPPSPPTPGSTLTTPPAVGLVHSG 516  
 RESULT 10  
 AAR22018  
 ID AAR22018 standard; Protein; 516 AA.  
 AC AAR22018;  
 XX 03-JUL-1992 (first entry)  
 DE Human thrombomodulin (1-516) with Asp367 substituted by Glu.  
 KW Mutant; thrombin binding site; blood clotting; TMm2 mutator.  
 OS Homo sapiens.  
 PN EP474273-A.  
 PD 11-MAR-1992.  
 XX 05-AUG-1991; 91EP-0202009.  
 PR 03-AUG-1990; 90JP-0204978.  
 XX (ASAH ) ASAH KASEI KOGYO.  
 PI Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;  
 DR WPI; 1992-081820/11.  
 XX New polypeptide inhibits blood coagulation and platelet  
 PT aggregation - promotes thrombin catalysed protein C activation  
 PT for treating myocardial infarction, thrombosis, embolism, etc.  
 XX Example 1; Page 18; 112pp; English.  
 PS Plasmid M13PMD1 (see AAR22016) encoding the first 516 N-terminal  
 CC amino acids of human thrombomodulin was used as a template for  
 CC site-directed mutagenesis using the "TMm2 mutator" to produce  
 CC plasmid M13TM2. In the mutant thrombomodulin encoded by the  
 CC plasmid, the wild-type Asp residue at position 367 is substituted  
 CC by an Glu residue. The activity of this mutant, truncated  
 CC thrombomodulin (i.e. D123Glu) in the activation of protein C was  
 CC compared to that of similarly truncated thrombomodulin with Asp  
 CC at position 367 (i.e. D123Asp). The activity of D123Glu was as high  
 CC as 1.2 times that of D123Asp. See AAR22013-R22022 and AAQ25074.  
 XX Sequence 516 AA;  
 SQ Query Match 99.9%; Score 2912; DB 13; Length 516;  
 Best Local Similarity 99.8%; Pred. No. 6.8e-162;

|||||  
Db 361 ECVEPVPDFRANCEYQCQLNLSILCVCAEGFAPIPHEPHRCQMFNCNACPADCDPN 420  
QY 421 TQASCECEGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDLSALVRHIGTDC 480  
Db 421 TQASCECEGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDLSALVRHIGTDC 480  
QY 481 DSGKVDGDSGSGEPPSPPTGTLTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPSPPTGTLTPPAVGLVHSG 516  
RESULT 8  
AAE23026  
ID AAE23026 standard; Protein; 575 AA.  
XX  
AC AAE23026;  
XX  
DT 21-AUG-2002 (first entry)  
XX  
DE Human thrombomodulin protein.  
XX  
KW Human; thrombomodulin; analogue; neurologic trauma; vulnerary;  
KW spinal cord injury; neurological damage; therapy; tranquilliser;  
KW neuroprotective.  
XX  
OS Homo sapiens.  
XX  
PH Key Location/Qualifiers  
FT Peptide 1..18  
FT Protein 19..575 /label= "Signal\_peptide"  
FT Domain 1..226 /note= "Human mature thrombomodulin"  
FT Domain 227..462 /note= "N-terminal domain"  
FT Modified-site 463..497 /note= "6 EGF-like domain"  
FT Domain 498..521 /note= "O-linked glycosylation site"  
FT Domain 522..557 /note= "Transmembrane domain"  
FT Domain /note= "Cytoplasmic domain"  
PN WO200217953-A2.  
XX  
PD 07-MAR-2002.  
XX  
PF 31-AUG-2001; 2001WO-US41930.  
XX  
PR 31-AUG-2000; 2000US-229714P.  
XX 23-AUG-2001; 2001US-0938405.  
XX  
PA (SCHD ) SCHERING AG.  
XX  
PI Festoff BW, Morser MJ;  
XX  
DR WPI; 2002-415551/44.  
XX N-PSDB; AAD36901.  
DR  
PT Treating neurological damage resulting from spinal cord injury in  
PT humans by administering a soluble, recombinant thrombomodulin analog  
PT which is resistant to oxidation -  
XX  
XX  
PS Claim 1; Fig 1; 31pp; English.  
XX  
CC The invention relates to a method of using thrombomodulin analogues  
CC in the treatment of the neurologic trauma associated with spinal  
CC cord injury in mammals. The method involves administering an oxidation  
CC resistant, soluble, recombinant thrombomodulin analogue, where the  
CC methionine at position 388 has been replaced with a leucine, and the  
CC analogue is numbered in accordance with native thrombomodulin. The  
CC method and thrombomodulin analogue are useful for treating

CC neurological damage resulting from spinal cord injury in human.  
CC The present sequence is human thrombomodulin protein.  
XX  
SQ Sequence 575 AA;  
Query Match 100.0%; Score 2916; DB 23; Length 575;  
Best Local Similarity 100.0%; Pred. No. 4.4e-162;  
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLGVLVLGALALAGLGFPAEPPQPGSQVEHDCFALYPGPATFLNASQICDGLRHLM 60  
Db 1 MLGVLVLGALALAGLGFPAEPPQPGSQVEHDCFALYPGPATFLNASQICDGLRHLM 60  
QY 61 TVRSSVAADVISLLINGDGGVRRRLWTGLQPLPGCGDPKRLGLRGQWTVGDNNTSYS 120  
Db 61 TVRSSVAADVISLLINGDGGVRRRLWTGLQPLPGCGDPKRLGLRGQWTVGDNNTSYS 120  
QY 121 RWARLDLNGAPLCPLCVAVSAAEATVPSEPIWEQQCEVADGFLCEFHFPATCRPLAV 180  
Db 121 RWARLDLNGAPLCPLCVAVSAAEATVPSEPIWEQQCEVADGFLCEFHFPATCRPLAV 180  
QY 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHAREAP 240  
Db 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVQGHAREAP 240  
QY 241 GAWDCSVENGGEHACNAIPGAPRCQCAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
Db 241 GAWDCSVENGGEHACNAIPGAPRCQCAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DPGSYSCMETGYRLAADOHRCEVDVDCILEPSPORCVNTQGGFECHECHYPNTDLVDG 360  
Db 301 DPGSYSCMETGYRLAADOHRCEVDVDCILEPSPORCVNTQGGFECHECHYPNTDLVDG 360  
QY 361 ECVEPVPDFRANCEYQCQLNLSILCVCAEGFAPIPHEPHRCQMFNCNACPADCDPN 420  
Db 361 ECVEPVPDFRANCEYQCQLNLSILCVCAEGFAPIPHEPHRCQMFNCNACPADCDPN 420  
QY 421 TQASCECEGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDLSALVRHIGTDC 480  
Db 421 TQASCECEGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDLSALVRHIGTDC 480  
QY 481 DSGKVDGDSGSGEPPSPPTGTLTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPSPPTGTLTPPAVGLVHSG 516  
RESULT 9  
AAE17521  
ID AAE17521 standard; Protein; 575 AA.  
XX  
AC AAE17521;  
XX  
DT 22-APR-2002 (first entry)  
XX  
DE Human full-length native thrombomodulin (TM).  
XX  
KW Human; thrombomodulin; TM analogue; protein C; thrombotic disease;  
KW antithrombotic therapy; thrombin-activatable fibrinolysis inhibitor;  
KW TAFI; myocardial infarction; disseminated intravascular coagulation;  
KW DIC; deep vein thrombosis; septic shock; pulmonary embolism; angina;  
KW acute respiratory distress syndrome; cancer; toxemia; septicemia;  
KW circulatory disorder; coronary embolism; pulmonary embolism; stroke;  
KW systemic coagulation disorder; immunosuppressive; cerebroprotective;  
KW pregnancy; anticoagulant; thrombolytic; venous occlusive condition;  
KW antibacterial; cardiant.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..18  
FT Protein 19..575 /label= "Signal\_peptide"  
FT /note= "Human mature thrombomodulin"

```

PR 30-JUL-1991; 91JP-0189984.
XX (ASAH ) ASahi CHEM IND CO LTD.
XX WPI; 1993-299652/38.
XX Novel polypeptide obtd. by culturing transformed fungus - having
PT blood coagulation preventing, platelet aggregation preventing and
PT thrombolytic activities
XX Disclosure; Fig 55; 65pp; Japanese.
PS Homo sapiens.
XX WO9322447-A.
XX 11-NOV-1993.
XX 30-APR-1993; 93WO-JP00578.
XX 01-MAY-1992; 92JP-0112903.
XX (ASAH ) ASahi CHEM IND CO LTD.
XX (ASAH ) ASahi KASEI KOGYO KK.
XX Kondo S, Toma K, Zushi M;
XX WPI; 1993-368806/46.
XX Peptide with anticoagulant and platelet aggregation inhibitor
PT activity - which promotes protein C activation by thrombin and is
PT useful in treating coagulation disorders e.g. thrombosis
XX Disclosure; Fig 1; 84pp; Japanese.
XX New peptides (see AARS0069) are inhibitors of the blood coagulation
CC and platelet aggregation activities of thrombin and promote the
CC protein-C activation effect of thrombin. They can be produced
CC efficiently in pure form by culture of appropriate transformants,
CC and are useful in treatment of circulatory disorders such as
CC myocardial infarction, thrombosis, embolism, telangiectasis,
CC arteriosclerosis obliterans, disseminated intravascular
CC coagulation, angina pectoris, gestosis and transient ischaemic
CC attack.
XX SQ Sequence 575 AA;
Query Match 100.0%; Score 2916; DB 14; Length 575;
Best Local Similarity 100.0%; Pred. No. 4.4e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLGVLVGLALAGLGPAPAEPPQPGSQCVHEHDFALYFGPATFLNASQICDGLRHLM 60
Db 1 MLGVLVGLALAGLGPAPAEPPQPGSQCVHEHDFALYFGPATFLNASQICDGLRHLM 60
QY 61 TVSSVAADVISILLNGDGGVRRRLWIGLQPPCGDPRKRLGRLGFQWTTGDNNTSYS 120
Db 61 TVSSVAADVISILLNGDGGVRRRLWIGLQPPCGDPRKRLGRLGFQWTTGDNNTSYS 120
QY 121 RWARLDNGAPLCGPLCAVSAEAATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAY 180
Db 121 RWARLDNGAPLCGPLCAVSAEAATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAY 180
QY 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPVAVQGHWAREAP 240
Db 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPVAVQGHWAREAP 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCETGYRLAADOHRCEVDVDCILEPSPQRCVNTQGGFECYCHYNYDLVDG 360
Db 301 DQPGSYSCMCETGYRLAADOHRCEVDVDCILEPSPQRCVNTQGGFECYCHYNYDLVDG 360
QY 361 ECVEPVDPCFRANCEYQCQPLNNTSYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPN 420
Db 361 ECVEPVDPCFRANCEYQCQPLNNTSYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPN 420
QY 421 TQASCECPGEGYILDDGICTDDCEBNGFCSCVCHNLPGTFFCICGPDGSAVVRHIGTDC 480
Db 421 TQASCECPGEGYILDDGICTDDCEBNGFCSCVCHNLPGTFFCICGPDGSAVVRHIGTDC 480
QY 481 DSGKVDGSGSGGEPPTPGSTLTPPAVGLVHSG 516
Db 481 DSGKVDGSGSGGEPPTPGSTLTPPAVGLVHSG 516
RESULT 7
AAR43031
XX AAR43031 standard; protein; 575 AA.
XX AC AAR43031;
XX DT 16-MAY-1994 (first entry)
XX

```

```

DE Human thrombomodulin.
XX Anticoagulant; platelet aggregation inhibitor;
XX protein C; activation; thrombin; thrombomodulin;
KW coagulation disorder; thrombosis; myocardial infarction;
KW embolism; telangiectasis; arteriosclerosis obliterans;
KW disseminated intravascular coagulation; DIC; angina pectoris;
KW gestosis; transient ischaemic attack.
XX OS Homo sapiens.
XX WO9322447-A.
XX 11-NOV-1993.
XX 30-APR-1993; 93WO-JP00578.
XX 01-MAY-1992; 92JP-0112903.
XX (ASAH ) ASahi CHEM IND CO LTD.
XX (ASAH ) ASahi KASEI KOGYO KK.
XX Kondo S, Toma K, Zushi M;
XX WPI; 1993-368806/46.
XX Peptide with anticoagulant and platelet aggregation inhibitor
PT activity - which promotes protein C activation by thrombin and is
PT useful in treating coagulation disorders e.g. thrombosis
XX Disclosure; Fig 1; 84pp; Japanese.
XX New peptides (see AARS0069) are inhibitors of the blood coagulation
CC and platelet aggregation activities of thrombin and promote the
CC protein-C activation effect of thrombin. They can be produced
CC efficiently in pure form by culture of appropriate transformants,
CC and are useful in treatment of circulatory disorders such as
CC myocardial infarction, thrombosis, embolism, telangiectasis,
CC arteriosclerosis obliterans, disseminated intravascular
CC coagulation, angina pectoris, gestosis and transient ischaemic
CC attack.
XX SQ Sequence 575 AA;
Query Match 100.0%; Score 2916; DB 14; Length 575;
Best Local Similarity 100.0%; Pred. No. 4.4e-162;
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MLGVLVGLALAGLGPAPAEPPQPGSQCVHEHDFALYFGPATFLNASQICDGLRHLM 60
Db 1 MLGVLVGLALAGLGPAPAEPPQPGSQCVHEHDFALYFGPATFLNASQICDGLRHLM 60
QY 61 TVSSVAADVISILLNGDGGVRRRLWIGLQPPCGDPRKRLGRLGFQWTTGDNNTSYS 120
Db 61 TVSSVAADVISILLNGDGGVRRRLWIGLQPPCGDPRKRLGRLGFQWTTGDNNTSYS 120
QY 121 RWARLDNGAPLCGPLCAVSAEAATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAY 180
Db 121 RWARLDNGAPLCGPLCAVSAEAATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAY 180
QY 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPVAVQGHWAREAP 240
Db 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPVAVQGHWAREAP 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCETGYRLAADOHRCEVDVDCILEPSPQRCVNTQGGFECYCHYNYDLVDG 360
Db 301 DQPGSYSCMCETGYRLAADOHRCEVDVDCILEPSPQRCVNTQGGFECYCHYNYDLVDG 360
QY 361 ECVEPVDPCFRANCEYQCQPLNNTSYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPN 420
Db 361 ECVEPVDPCFRANCEYQCQPLNNTSYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPN 420

```

QY 121 RWARLDLNGAPLCGLCVAVSAAEATVPSEPIWEEOQCEVKADGFLCEHFHPATCRPLAV 180  
DB 121 RWARLDLNGAPLCGLCVAVSAAEATVPSEPIWEEOQCEVKADGFLCEHFHPATCRPLAV 180  
QY 181 EPGAAAAVSIYTGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
DB 181 EPGAAAAVSIYTGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
QY 241 GAWDCSVENGGEHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
DB 241 GAWDCSVENGGEHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360  
DB 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360  
QY 361 ECVEPVDFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFCTACPADCDPN 420  
DB 361 ECVEPVDFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFCTACPADCDPN 420  
QY 421 TQASCECPGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDPSALVRHIGTDC 480  
DB 421 TQASCECPGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDPSALVRHIGTDC 480  
QY 481 DSGKVDGSDSGSGEPPTPGSTLTTPPAVGLVHSG 516  
DB 481 DSGKVDGSDSGSGEPPTPGSTLTTPPAVGLVHSG 516  
RESULT 5  
AAR11534  
ID AAR11534 standard; Protein; 575 AA.  
XX AAR11534;  
XX AC AAR11534;  
XX FT Peptide 1..18 Location/Qualifiers  
XX FT Peptide /label= signal peptide  
XX FT Peptide 19..365  
XX FT Peptide /label= opt. N-terminal sequence  
XX FT Peptide 366..480  
XX FT Peptide /label= human thrombomodulin active fragment  
XX FT Peptide 481..516  
XX FT Peptide /label= opt. C-terminal sequence  
XX PN WO9104276-A.  
XX XX  
XX PD 04-APR-1991.  
XX XX  
XX PF 25-SEP-1990; 90WO-JP01234.  
XX XX  
XX PR 25-SEP-1989; 85JP-0246270.  
XX XX  
XX PA (ASAH ) ASAH KASEI KOGYO.  
XX PI Yamamoto S, Gomi K, Ogawa K;  
XX XX  
XX DR WPI; 1991-117478/16.  
XX XX  
XX PT New human thrombomodulin polypeptide contg. chondroitin - has  
XX PT anticoagulant platelet aggregation and thrombolytic activity  
XX PS  
XX PS Disclosure; fig 1; 80pp; Japanese.  
XX XX  
XX CC This human thrombomodulin polypeptide comprises a 115 residue active  
XX CC fragment with opt. flanking sequences (N-terminal= 347 amino acids;

CC C-terminal= 36 amino acids). Additionally it has an N-terminal signal  
CC sequence and an attached sugar chain, contg. chondroitin or chondroitin  
CC sulphate. It promotes the activation of protein C by thrombin and has  
CC good anticoagulant, platelet aggregation inhibition and thrombolytic  
CC activities. It can therefore be used for treating cardiovascular  
CC diseases, eg arteriosclerosis or myocardial infarction.  
XX  
SQ Sequence 575 AA;  
Query Match 100.0%; Score 2916; DB 12; Length 575;  
Best Local Similarity 100.0%; Pred. No. 4.4e-162;  
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MLGVVLGALALAGLPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRHLM 60  
DB 1 MLGVVLGALALAGLPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRHLM 60  
QY 61 TVRSVAADVLSILLNGDGGVRRRLWTGLQPPCGDPPKRLPLGRGQWVTGDNNTSYS 120  
DB 61 TVRSVAADVLSILLNGDGGVRRRLWTGLQPPCGDPPKRLPLGRGQWVTGDNNTSYS 120  
QY 121 RWARLDLNGAPLCGLCVAVSAAEATVPSEPIWEEOQCEVKADGFLCEHFHPATCRPLAV 180  
DB 121 RWARLDLNGAPLCGLCVAVSAAEATVPSEPIWEEOQCEVKADGFLCEHFHPATCRPLAV 180  
QY 181 EPGAAAAVSIYTGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
DB 181 EPGAAAAVSIYTGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
QY 241 GAWDCSVENGGEHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
DB 241 GAWDCSVENGGEHACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360  
DB 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360  
QY 361 ECVEPVDFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFCTACPADCDPN 420  
DB 361 ECVEPVDFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFCTACPADCDPN 420  
QY 421 TQASCECPGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDPSALVRHIGTDC 480  
DB 421 TQASCECPGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDPSALVRHIGTDC 480  
QY 481 DSGKVDGSDSGSGEPPTPGSTLTTPPAVGLVHSG 516  
DB 481 DSGKVDGSDSGSGEPPTPGSTLTTPPAVGLVHSG 516  
RESULT 6  
AAR41806  
ID AAR41806 standard; peptide; 575 AA.  
XX  
XX AC AAR41806;  
XX XX  
XX DT 30-MAR-1994 (first entry)  
XX XX  
XX DE Thrombomodulin.  
XX KW Transformation; fungus; blood coagulation; prevention; platelet;  
XX KW aggregation; thrombolytic activity; thrombolysis; human thrombomodulin;  
XX KW site-directed mutagenesis; promotion; protein C; activation; thrombin.  
XX OS Homo sapiens.  
XX XX  
XX PN JP05213998-A.  
XX XX  
XX PD 24-AUG-1993.  
XX XX  
XX PF 02-AUG-1991; 91JP-0282369.  
XX XX  
XX PR 03-AUG-1990; 90JP-0204978.

QY 421 TQASCEPEGYILDDGFICTDIDECENGFGSGVCHNLPGTFECICGPDPSALVRHIGTDC 480  
Db 421 TQASCEPEGYILDDGFICTDIDECENGFGSGVCHNLPGTFECICGPDPSALVRHIGTDC 480  
QY 481 DSGKVDGDSGSGEPPTGCTLTTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPTGCTLTTPPAVGLVHSG 516

RESULT 3  
AAV83934  
ID AAV83934 standard; Protein; 516 AA.  
XX  
AC AAV83934;  
XX  
DT 28-JUL-2000 (first entry)  
XX  
DE Human thrombomodulin TMD protein.  
XX  
KW Human; thrombomodulin; vasculitis; protein C; thrombin.  
XX  
OS Homo sapiens.  
XX  
PN JP2000053582-A.  
XX  
PD 22-FEB-2000.  
XX  
PF 06-AUG-1998; 98JP-0222688.  
XX  
PR 06-AUG-1998; 98JP-0222688.  
XX  
PA (ASAH ) ASAH KASEI KOGYO KK.  
XX  
PA WPI; 2000-353249/31.  
DR N-PSDB; AAA10027.  
XX  
PT Treating agent for vasculitis contains peptide which promotes  
activation of protein C by thrombin -  
XX  
PS Claim 4; Page 10-12; 18pp; Japanese.  
XX  
CC This sequence represents a human thrombomodulin protein. The invention  
relates to a treating agent for vasculitis containing a peptide which  
promotes activation of protein C by thrombin.  
XX  
SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 21; Length 516;  
Best Local Similarity 100.0%; Pred. No. 4e-162;  
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVLVLGALALAGLGFAPAEPQPGSQCVQVEHDCFALYPCPATFLNASQICDGLRGHLM 60  
Db 1 MGVLVLGALALAGLGFAPAEPQPGSQCVQVEHDCFALYPCPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQPLPGCGDPKRLGRLGFWVTGDNNTSYS 120  
Db 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQPLPGCGDPKRLGRLGFWVTGDNNTSYS 120  
QY 121 RWARLDNGAPLCGLPVAVSAEATVPSEPIWEBOQCEVKADGFLCFEFPATCRPLAV 180  
Db 121 RWARLDNGAPLCGLPVAVSAEATVPSEPIWEBOQCEVKADGFLCFEFPATCRPLAV 180  
QY 181 EFGAAAVASITYGTFPAARGADFOALPVGSSAAVAPLGLQIMCTAPPAGVGHWAREAP 240  
Db 181 EFGAAAVASITYGTFPAARGADFOALPVGSSAAVAPLGLQIMCTAPPAGVGHWAREAP 240  
QY 241 GAWDCSVENGCGEACNAIPGAPRCQCPAGALQADGSRCTASATQSCNDLCEHFCVNP 300  
Db 241 GAWDCSVENGCGEACNAIPGAPRCQCPAGALQADGSRCTASATQSCNDLCEHFCVNP 300  
QY 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILPSPQRCVNTQGGFECHECHYPNYDLVDG 360  
Db 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILPSPQRCVNTQGGFECHECHYPNYDLVDG 360

Db 301 DPGSYSCMCETGYRLAADQHRCEVDVDCILPSPQRCVNTQGGFECHECHYPNYDLVDG 360  
QY 361 ECVEPVDPCFRANCEYOCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNQTACPADCDPN 420  
Db 361 ECVEPVDPCFRANCEYOCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFNQTACPADCDPN 420  
QY 421 TQASCEPEGYILDDGFICTDIDECENGFGSGVCHNLPGTFECICGPDPSALVRHIGTDC 480  
Db 421 TQASCEPEGYILDDGFICTDIDECENGFGSGVCHNLPGTFECICGPDPSALVRHIGTDC 480  
QY 481 DSGKVDGDSGSGEPPTGCTLTTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPTGCTLTTPPAVGLVHSG 516

RESULT 4  
AAV69529  
ID AAV69529 standard; Protein; 516 AA.  
XX  
AC AAV69529;  
XX  
DT 10-APR-2000 (first entry)  
XX  
DE Human thrombomodulin variant, SEQ ID NO:3.  
XX  
KW Thrombomodulin; TM; soluble; affinity purification; cation exchange;  
anticoagulant; thrombolytic.  
XX  
OS Homo sapiens.  
XX  
PN JP11341990-A.  
XX  
PD 14-DEC-1999.  
XX  
PF 23-MAR-1999; 99JP-0077518.  
XX  
PR 30-MAR-1998; 98JP-0084389.  
XX  
PA (ASAH ) ASAH KASEI KOGYO KK.  
XX  
PA WPI; 2000-101696/09.  
DR N-PSDB; AAZ55965.  
XX  
PT Preparation of highly pure soluble thrombomodulin - used as an  
antiblood coagulation agent and a thrombolytic agent  
XX  
PS Claim 9; Page 30-32; 36pp; Japanese.  
XX  
CC The invention relates to a novel method for the preparation of highly  
pure soluble thrombomodulin (TM) containing substantially no serum-  
derived or antibody-derived substance. The method comprises isolating  
soluble TM from unpurified serum via affinity purification using an  
anti-TM antibody. The soluble TM is then treated with a cation  
exchanger at a specific conductivity of 25 to 34 ms/cm and a pH of 3 to  
4, and the fraction containing the soluble TM is isolated. The soluble  
TM can be used as a blood anticoagulation agent and a thrombolytic agent.  
XX  
SQ This sequence represents a human thrombomodulin variant, designated  
SEQ ID NO:3.

Query Match 100.0%; Score 2916; DB 21; Length 516;  
Best Local Similarity 100.0%; Pred. No. 4e-162;  
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVLVLGALALAGLGFAPAEPQPGSQCVQVEHDCFALYPCPATFLNASQICDGLRGHLM 60  
Db 1 MGVLVLGALALAGLGFAPAEPQPGSQCVQVEHDCFALYPCPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQPLPGCGDPKRLGRLGFWVTGDNNTSYS 120  
Db 61 TVRSSVAADVISLLNGDGGVGRRLRWIGLQPLPGCGDPKRLGRLGFWVTGDNNTSYS 120



XX Plasmid M13mp19TMJ3 (see AAR22013) was used as template for  
CC site-directed mutagenesis. A 177bp fragment was deleted using the  
CC "rmd1 deleter" oligonucleotide to give plasmid M13TMD1 which encodes  
CC the first 516 N-terminal amino acids of human thrombomodulin. Plasmid  
CC M13TMD1 was completely digested with HindIII and BamHI and a TMD1  
CC -fragment of ca. 1700bp was isolated. The fragment was ligated to  
CC HindIII- and BglII-cut plasmid pSV2-dhfr to give the recombinant  
CC plasmid pSV2TMD1. See AAR22014-R22022 and AAQ25072.  
XX  
SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 13; Length 516;  
Best Local Similarity 100.0%; Pred. No. 4e-162;  
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVVLGALALAGLGPAPAEPPGSGQCVHDCFCALYPGATFLNASQICDGLRGHLM 60  
DB 1 MGVVLGALALAGLGPAPAEPPGSGQCVHDCFCALYPGATFLNASQICDGLRGHLM 60

QY 61 TVRSSVAADVVISLLNGDGGVRRRLWIGLQPPCGDPRKRLGFRQWVTGDNNTSYS 120  
DB 61 TVRSSVAADVVISLLNGDGGVRRRLWIGLQPPCGDPRKRLGFRQWVTGDNNTSYS 120

QY 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
DB 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180

QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
DB 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240

QY 241 GAWDCSVENGCGEHACNAIPCAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
DB 241 GAWDCSVENGCGEHACNAIPCAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300

QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
DB 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360

QY 421 TSACECEGYLDDGFTCTDIDECENGFCGVCNHLPGTFCICGPDLSALVRHIGTDC 480  
DB 421 TSACECEGYLDDGFTCTDIDECENGFCGVCNHLPGTFCICGPDLSALVRHIGTDC 480

QY 481 DSGKVDGSDSGSEPPSPPTPGSTILTPPAVGLVHSG 516  
DB 481 DSGKVDGSDSGSEPPSPPTPGSTILTPPAVGLVHSG 516

RESULT 2  
AA09347  
ID AA09347 standard; Protein; 516 AA.  
XX  
AC AA09347;  
XX  
DT 08-JUL-1999 (first entry)  
XX  
DE Human thrombomodulin SEQ ID NO:1.  
XX  
KW Human; thrombomodulin; aqueous parenteral solution; storage;  
KW distribution; acute coronary syndrome; thrombosis; embolism;  
KW diabetes.  
XX  
OS Homo sapiens.  
XX  
PN WO9918994-A1.  
XX  
PD 22-APR-1999.  
XX

PF 13-OCT-1998; 98WO-JP04609.  
XX  
PR 11-NOV-1997; 97JP-0308523.  
PR 15-OCT-1997; 97JP-0281659.  
XX  
PA (ASAH ) ASahi KASEI KOGYO KK.  
XX  
PI Murata T, Shimizu H, Tsuruta K, Yokozawa A, Yui M;  
XX  
DR WPT; 1999-277444/23.  
DR N-PSDB; AAX55879.  
XX  
PT Stable aqueous parenteral thrombomodulin solution - comprising  
PT buffer and surfactant, useful for treating acute coronary syndrome,  
PT thrombosis, embolism, and diabetes  
XX  
PS Claim 6; Page 87-89; 97pp; Japanese.  
XX  
CC The present invention describes a method for maintaining the quality of  
CC an aqueous parenteral solution of thrombomodulin comprising buffer and  
CC surfactant aseptically filled in a case or syringe. Maintaining the  
CC quality of an aqueous, parenteral thrombomodulin solution is  
CC characterised in that the solution: (a) comprises soluble thrombomodulin,  
CC a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically  
CC filled into: (i) a case or (ii) a syringe without any empty space; and  
CC (c) is kept in liquid form in storage and distribution and not frozen or  
CC freeze dried. Thrombomodulin is used to treat acute coronary syndrome  
CC (such as myocardial infarction, unstable angina and coronary artery  
CC blockage), thrombosis (e.g. cerebral, vascular and peripheral blood  
CC vessel thrombosis), embolism, peripheral blood vessel disorders (e.g.  
CC Raynaud's disease), arteriosclerosis, vasculitis (e.g. due to systemic  
CC lupus erythematosus or Barrett's syndrome), multiple organ failure,  
CC disseminated intravascular coagulation, transient ischaemia, diabetes,  
CC liver veno-occlusive diseases and deep vein thrombosis. The composition  
CC is stable for a long period of time and can be stored and distributed in  
CC ready to use form avoiding the problems of dissolution and accuracy when  
CC preparing on demand. The present sequence represents human  
CC thrombomodulin.  
XX  
SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 20; Length 516;  
Best Local Similarity 100.0%; Pred. No. 4e-162;  
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGVVLGALALAGLGPAPAEPPGSGQCVHDCFCALYPGATFLNASQICDGLRGHLM 60  
DB 1 MGVVLGALALAGLGPAPAEPPGSGQCVHDCFCALYPGATFLNASQICDGLRGHLM 60

QY 61 TVRSSVAADVVISLLNGDGGVRRRLWIGLQPPCGDPRKRLGFRQWVTGDNNTSYS 120  
DB 61 TVRSSVAADVVISLLNGDGGVRRRLWIGLQPPCGDPRKRLGFRQWVTGDNNTSYS 120

QY 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
DB 121 RWARLDLNGAPLCGPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180

QY 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
DB 181 EPGAAAASVITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240

QY 241 GAWDCSVENGCGEHACNAIPCAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
DB 241 GAWDCSVENGCGEHACNAIPCAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300

QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
DB 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360

QY 361 ECVEPVDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNACPADCPN 420  
DB 361 ECVEPVDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNACPADCPN 420

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:19:38 ; Search time 34.5 Seconds  
(without alignments)  
1992.966 Million cell updates/sec

**Title:** US-09-509-994-1

Perfect score:

Sequence: 1 MLGVLVLGALALAGLGFAP.....PSPTPGSTLTTPPAVGLVHSG 516

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database': A\_Geneseq 101002:★

- 1: /SID52/gcgdata/genseq/genseqp-emb1/AA1980.DAT:\*
- 2: /SID52/gcgdata/genseq/genseqp-emb1/AA1981.DAT:\*
- 3: /SID52/gcgdata/genseq/genseqp-emb1/AA1982.DAT:\*
- 4: /SID52/gcgdata/genseq/genseqp-emb1/AA1983.DAT:\*
- 5: /SID52/gcgdata/genseq/genseqp-emb1/AA1984.DAT:\*
- 6: /SID52/gcgdata/genseq/genseqp-emb1/AA1985.DAT:\*
- 7: /SID52/gcgdata/genseq/genseqp-emb1/AA1986.DAT:\*
- 8: /SID52/gcgdata/genseq/genseqp-emb1/AA1987.DAT:\*
- 9: /SID52/gcgdata/genseq/genseqp-emb1/AA1988.DAT:\*
- 10: /SID52/gcgdata/genseq/genseqp-emb1/AA1989.DAT:\*
- 11: /SID52/gcgdata/genseq/genseqp-emb1/AA1990.DAT:\*
- 12: /SID52/gcgdata/genseq/genseqp-emb1/AA1991.DAT:\*
- 13: /SID52/gcgdata/genseq/genseqp-emb1/AA1992.DAT:\*
- 14: /SID52/gcgdata/genseq/genseqp-emb1/AA1993.DAT:\*
- 15: /SID52/gcgdata/genseq/genseqp-emb1/AA1994.DAT:\*
- 16: /SID52/gcgdata/genseq/genseqp-emb1/AA1995.DAT:\*
- 17: /SID52/gcgdata/genseq/genseqp-emb1/AA1996.DAT:\*
- 18: /SID52/gcgdata/genseq/genseqp-emb1/AA1997.DAT:\*
- 19: /SID52/gcgdata/genseq/genseqp-emb1/AA1998.DAT:\*
- 20: /SID52/gcgdata/genseq/genseqp-emb1/AA1999.DAT:\*
- 21: /SID52/gcgdata/genseq/genseqp-emb1/AA2000.DAT:\*
- 22: /SID52/gcgdata/genseq/genseqp-emb1/AA2001.DAT:\*
- 23: /SID52/gcgdata/genseq/genseqp-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query %		Length	DB	ID	Description
		Match					
1	2916	100.0	516	13	AAR22016		Truncated human th
2	2916	100.0	516	20	AAY09347		Human thrombomod
3	2916	100.0	516	21	AAY83934		Human thrombomod
4	2916	100.0	516	21	AAY69529		Human thrombomod
5	2916	100.0	575	12	AAR11534		Human thrombomod
6	2916	100.0	575	14	AAR41806		Thrombomodulin, H
7	2916	100.0	575	14	AAR43031		Human thrombomod
8	2916	100.0	575	23	AAE23026		Human thrombomod
9	2916	100.0	575	23	AAE17521		Human full-length
10	2912	99.9	516	13	AAR22016		Human thrombomod

11	2912	99.9	516	20	AA093348	Human thrombomodul
12	2912	99.9	516	21	AA093348	Human thrombomodul
13	2908	99.7	516	13	AA022017	Human thrombomodul
14	2906	99.7	515	12	AA010617	Soluble thrombomod
15	2906	99.7	575	9	AA082070	Human thrombomodul
16	2904	99.6	575	14	AA031572	Human thrombomodul
17	2899	99.4	575	13	AA020639	Human urinary thro
18	2875	98.6	575	20	AA073970	Human thrombomodul
19	2858	98.0	575	13	AA022189	Sequence of thromb
20	2830	97.1	498	16	AA084185	Human derived thro
21	2830	97.1	498	18	AA001600	Thrombomodulin TME
22	2827	96.9	557	23	AA023028	Human thrombomodul
23	2827	96.9	557	23	AA023031	Human thrombomodul
24	2827	96.9	557	23	AA017525	Human thrombomodul
25	2827	96.9	557	23	AA017528	Human thrombomodul
26	2826	96.9	557	23	AA017523	Human thrombomodul
27	2825	96.9	498	21	AA067402	Human thrombomodul
28	2825	96.9	557	23	AA017596	Novel sugar chain-
29	2824	96.8	497	19	AA069520	Human thrombomodul
30	2824	96.8	557	23	AA017530	rs7M protein SEQ I
31	2824	96.8	557	23	AA017538	Human thrombomodul
32	2823	96.8	557	23	AA023029	Human thrombomodul
33	2823	96.8	557	23	AA017526	Human thrombomodul
34	2822	96.8	557	23	AA023030	Human thrombomodul
35	2822	96.8	557	23	AA017537	Human thrombomodul
36	2822	96.8	557	23	AA017529	Human thrombomodul
37	2822	96.8	557	23	AA017593	Human thrombomodul
38	2822	96.8	557	23	AA017595	Human thrombomodul
39	2822	96.8	557	23	AA017597	Human thrombomodul
40	2821	96.7	557	23	AA017591	Human thrombomodul
41	2821	96.7	557	23	AA017592	Human thrombomodul
42	2821	96.7	557	23	AA017594	Human thrombomodul
43	2820	96.7	497	17	AA094607	Human recombinant
44	2820	96.7	557	23	AA017524	Human thrombomodul
45	2815	96.5	554	23	AA023027	Human thrombomodul

## ALIGNMENTS

RESULT 1	
AAAR22016	
ID	AAAR22016 standard; Protein; 516 AA.
XX	
AC	AAAR22016;
XX	
DT	03-JUL-1992 (first entry)
XX	
DE	Truncated human thrombomodulin encoded by plasmid pSV2TMD1.
XX	
XX	
KW	Thrombin binding site; blood clotting; TMD1 deleter.
XX	
OS	Homo sapiens.
XX	
PN	EP474273-A.
XX	
PD	11-MAR-1992.
XX	
PF	05-AUG-1991; 91EP-0202009.
XX	
PR	03-AUG-1990; 90JP-0204978.
XX	
PA	(ASAH ) ASahi Kasei Kogyo.
XX	
PI	Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;
XX	
DR	WPI; 1992-081820/11.
XX	
PT	New polypeptide inhibits blood coagulation and platelet
PT	aggregation - promotes thrombin catalysed protein C activat
PT	for treating myocardial infarction, thrombosis, embolism, e
XX	
PS	Example 1; Page 17; 112pp; English.

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:26:23 ; Search time 8.5 Seconds  
(without alignments)  
1011.502 Million cell updates/sec

Title: US-09-509-994-1  
Perfect score: 2916  
Sequence: 1 MLGVLVGLALAGLGFAP.....PSPTPGSTLPPAVGLVHSG 516

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105981 seqs, 16662342 residues

Total number of hits satisfying chosen parameters: 105981

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2916	100.0	575	10	US-09-938-405-2
2	581	19.9	652	10	US-09-789-919-96
3	543	18.6	644	10	US-09-789-919-62
4	543	18.6	644	10	US-09-789-919-73
5	355	12.2	776	9	US-10-000-512-8
6	355	12.2	959	9	US-10-000-512-10
7	352.5	12.1	877	10	US-09-764-898-200
8	350.5	12.0	915	9	US-09-905-291A-34
9	350.5	12.0	915	10	US-09-909-320-34
10	350.5	12.0	915	10	US-09-909-088B-34
11	333	11.4	999	10	US-09-747-371-2
12	328.5	11.3	1009	10	US-09-898-570-16
13	327.5	11.2	974	10	US-09-898-570-14
14	326	11.2	800	10	US-09-764-853-800
15	326	11.2	800	10	US-09-764-898-272
16	323	11.1	534	10	US-09-804-156-14
17	323	11.1	534	10	US-09-946-633-6
18	315.5	10.8	845	10	US-09-898-570-12
19	315	10.8	997	10	US-09-747-371-3

ALIGNMENTS

RESULT 1

US-09-938-405-2

; Sequence 2, Application US/09938405

; Patent No. US20020111296A1

; GENERAL INFORMATION:

; APPLICANT: Restoff, Barry W.

; APPLICANT: Morser, Michael J.

; TITLE OF INVENTION: Thrombomodulin Analogs for Use in Recovery of Spinal Cord Injury

; FILE REFERENCE: 51960AUSMI

; CURRENT APPLICATION NUMBER: US/09/938,405

; CURRENT FILING DATE: 2001-08-23

; PRIOR APPLICATION NUMBER: 60/229,714

; PRIOR FILING DATE: 2000-08-31

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 575

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-938-405-2

Query Match 100.0%; Score 2916; DB 10; Length 575;  
Best Local Similarity 100.0%; Pred No. 2,4e-174;  
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVGLALAGLGFAPPAEPQPGSCQVHDCFCALYPGPATFLNASQICDGLRGLHM 60

Db 1 MLGVLVGLALAGLGFAPPAEPQPGSCQVHDCFCALYPGPATFLNASQICDGLRGLHM 60

QY 61 TVRSSVAADVISLLNGDGVGRRRLWIGLQLPPGGDPRLGLRGFQWTDGNTSYS 120

Db 61 TVRSSVAADVISLLNGDGVGRRRLWIGLQLPPGGDPRLGLRGFQWTDGNTSYS 120

QY 121 RWARLDNGAPLGGPLCVASAAEATVPSEPIWEEQCEVKADGFLCEFFHPPATCRPLAV 180

Db 121 RWARLDNGAPLGGPLCVASAAEATVPSEPIWEEQCEVKADGFLCEFFHPPATCRPLAV 180

QY 181 EPGAAAAAIVSYTGTFFAARGADFQALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240

Db 181 EPGAAAAAIVSYTGTFFAARGADFQALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240

QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVPNP 300

```
Db 241 GAWDCSVENGCGEACNAIPGRCQCPAGAAQADGRSCTASATSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPCQRCVNTGGFCHCYPNYDLVDG 360
Db 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPCQRCVNTGGFCHCYPNYDLVDG 360
QY 361 ECVEPVPDPCFRANCEYQCPNLQNTSYLCVCAEGFAPIPHEPHRCQMCNQTACADCDPN 420
Db 361 ECVEPVPDPCFRANCEYQCPNLQNTSYLCVCAEGFAPIPHEPHRCQMCNQTACADCDPN 420
QY 421 TOASCEPEGYILDGFTCTDIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDC 480
Db 421 TOASCEPEGYILDGFTCTDIDECENGFCGSGVCHNLPGTFECICGPDLSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGPPSPPTPGSTLTTPPAVGLVHSG 516
Db 481 DSGKVDGDSGSGPPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 2
US-09-789-919-96
; Sequence 96, Application US/09789919
; Patent No. US20020064855A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor
; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM
; FILE REFERENCE: 2275-1-005
; CURRENT APPLICATION NUMBER: US/09/789,919
; CURRENT FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 96
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-789-919-96

Query Match 19.9%; Score 581; DB 10; Length 652;
Best Local Similarity 31.6%; Pred. No. 2.8e-29;
Matches 180; Conservative 56; Mismatches 210; Indels 124; Gaps 32;

QY 2 LGVLVGLAALA--GLGFAPAPBPQSGSOVHDCFALYPGPATFLNASQICDGLRGLH 59
Db 5 MGLLLLLLLLLTOPAGTGADTE-----AVVCGTACTYTAHSGKLSAAEAQHNCNQNGNL 60
QY 60 MTRVSSVAAD-----VISLLNGDGGVGR--RLWIGLQLPPG--CGDPKRLGRLGFOWVT 112
Db 61 ATVKSKEAQHVQVRLAQLLRREALTARMSKFWIGLQREKGLCLDPSL--PLKGFSWVG 118
QY 113 GDNNTSYRWARDLNGAPLCGLPCA--VSAAEATVPSE--PIWEEQOC-----EYKAD 163
Db 119 GGEDTPYSNNHKELRNSC--ISKRCVSLLDLSQPLLPNLPKWSGCPGSPGSGNIE 176
QY 164 GFICEFHFPATCRPLAV--EPGAAAASVITYTTPFAARGADFOALPVGSSAAVAPGLQL 222
Db 177 GFVCKFSFGMCRPLALGPG-----QVTTYTTPQTSSLEAVFPASAANVA----- 224
QY 223 MC-TAPPAGVQGHW---AREAPGAWD-----CSVENGGCEHAC--NAIPGAP 263
Db 225 -CGEGDKDETQSHYFLCKEKAPDFWNGSSGPLCVKPKYCNFNNGGCHODCFEGGDSF 283
QY 264 RCQCPAGAAQADGRSCTASATSCNDLCE--HFCVNPNDQPGSYSCMCTGYRLAADQH 321
Db 284 LCGRCPGRRLLDLVTC-ASRNPSCSSPCRGATCVLGP-HGKNYTCRCPPGQQLDSSQL 341
QY 322 RCEDVDVDCILEPSPCQRCVNTGGFCHCYPNYDLVDGECVBPDPFCFRANCEYQCPQL 381
Db 342 DCVDVDEC--QDSPCAQECVNTPGFRCECWGY-----EPGGP-----GEGACQDV 386
QY 382 NQTSYLCVCAEGFAPIPHEPHRCQMCNQTACADCDPNQAS--CECPGYIL--DDGF 437
```

```
Db 387 DE-----CALGRSP-----CAQGC-TNTDGSFHCSEEGYVLAGEGT 423
QY 438 ICTDIDEC--ENGFGCSGVCHNLPGTFECICGPDLSALVRHIGTDCDSKV-----D 486
Db 424 QCQDVDECVGPGGLCDLSLCFNQTQGSFHCGLPGVWLAPN-GVSCMTGMPVSLGPPSGPPD 482
QY 487 GGDGSGSGE-----PPSPPTPGSTLTTPPA 509
Db 483 EEDKGEKEGSTVPRAATASPTRGPEGTPKA 512

RESULT 3
US-09-789-919-62
; Sequence 62, Application US/09789919
; Patent No. US20020064855A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor
; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM
; FILE REFERENCE: 2275-1-005
; CURRENT APPLICATION NUMBER: US/09/789,919
; CURRENT FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-789-919-62

Query Match 18.68; Score 543; DB 10; Length 644;
Best Local Similarity 30.0%; Pred. No. 6.3e-27;
Matches 167; Conservative 55; Mismatches 211; Indels 124; Gaps 27;

QY 5 LVLGALALAGLGFAPAPBPQSGSOVHDCFALYPGPATFLNASQICDGLRGLHMTVRS 64
Db 9 LLLGLLQOPWAG--AAADSQ--AVVCGTACTYTAHSGKLSAAEAQHRCNENGNNLATVKS 64
QY 65 SVAA-----DVISLLNGDGGVGR--RLWIGLQLPPGCGDPKRLGRLGFOWVTGDNNTS 118
Db 65 EEEARHVQQAQLTQLLTKAPLEAKWGFWIGLQREKGNCTYHDL-PMRGFSWVGGEETA 123
QY 119 YSRWARDLNGAPLCGLPCA--VSAAEATVPSE--PIWEEQOC-----KADGFLCEPHF 171
Db 124 YSNWYKASKSSCIFKRCVSLLDLSLTPHPSHLPKWHESPCTPEAPGNSIEGLCKENF 183
QY 172 PATCRPLAV--EPGAAAASVITYTTPFAARGADFOALPVGSSAAVAPGLQ-----LM 223
Db 184 KGMRPLALGPG-----RVTTYTTPQTSSLEAVFPASVANVA--CGDEAKSETHYFL 236
QY 224 CT-APPAGVQGHWAREAP---GAWDCSVENGCGEAC--NAIPGAPRCQCPAGAAQADG 277
Db 237 CNEKTPGIF--HWGSSGPLCVSPKFGCSFNNGGQQDCFEFGDGSFRCGPRGFRLLDDL 294
QY 278 RCTASATQSCNDLCEHFCVNPNDQPG-----SYSCMCTGYRLAADQHRCEVDV 327
Db 295 VTCAS-----RNPCCSSNPCTGGMCHSVPLSENYTCRCPSGQQLDSSQVHCVDID 344
QY 328 DCILEPSPCQRCVNTGGFCHCYPNYDLVDG---ECVEPVPDPCFRANCEYQCPQLNQT 384
Db 345 EC--QDSPCAQDCVNTLGSFHCCEWGYQ-PSGPKKEACEVDVDECAANS------ 392
QY 385 SYLCVCAEGFAPIPHEPHRCQMCNQTACADCDPNQAS--CECPGYIL--DDGFICT 440
Db 393 -----CAQGC-----NDGFSFYCKEYIVSGEDSTQCE 423
QY 441 DIDECE--GGFGSGVCHNLPGTFECICGPDLSALVRHIGTDCDSG-----KVD 486
Db 424 DIDECSARNPCDLSLCFNQDGGFRCCPGWELAPN-GVFCSGTGVFSELPAAPPOKED 482
QY 487 GGDGSGSGPPSPPTPGS 503
```





[illegible]

## RESULT 9

```

US-09-909-320-34
; Sequence 34, Application US/09909320
; Patent No. US20020132240A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14*
; CURRENT APPLICATION NUMBER: US/09/909,320
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29

```

```

; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 34
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-909-320-34

Query Match      12.0%   Score 350.5;   DB 10;   Length 915;
Best Local Similarity 34.3%;   Pred. No. 8e-15;
Matches 89;   Conservative 37;   Mismatches 107;   Indels 19;   Gaps 14;

QY  245  CSVENGGCEHACNAPGAPRCQCPAGAAALQADGRSCTA-SATQSCNDLCEHFCVPNPDQP 303
DB  283  CAMEDHNCQEQLCVNPGSFVCQYSGYALADGKRCVADVYCASENHGCEHCY-NAD-- 339

QY  304  GSYSCMETGYRLAANDQHRCDVDVDCILEPSPQRCVNTGGFECHCYPNVDL-VDGEC 362
DB  340  GSYLCQCEHGFALNPDEKTCITRINYCALNKPCEHCEVCNMEESYCRCHRGYTLDPNGKT 399

QY  363  VEPVDPFCFRAN--CEYQCOPLN-QTSYLCVCAEGFAPIPHEPHRCQM--FC--NOTACPA 415
DB  400  CSRVVDHCAQDDHGCEQLC--LNTEDSFVCQCEGFL-INEDLKTCRSYVDYCLLSDHGCEY 456

QY  416  DC-DPNTQASCECPGEGYIL-DDGFGICTDIDSCENGGF--CSGVCHNLPTFCFCICGPD$AL 472
DB  457  SCVNMDRSFACQCPGEGHVLRSDDGKTCALKLDSCALGDHCCHEHSCVSSDSFVCQCF-EGYI 515

QY  473  VRHIGTDC 480
DB  516  LREDGKTC 523

RESULT 10
US-09-909-088B-34
; Sequence 34, Application US/09909088B
; Patent No. US20020146709A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Flivaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.

```

```

/ APPLICANT: Tunas, Daniel
/ APPLICANT: Williams, P. Mickey
/ APPLICANT: Wood, William, I.
/ TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
/ TITLE OF INVENTION: Acids Encoding the Same
/ FILE REFERENCE: 10466-14
/ CURRENT APPLICATION NUMBER: US/09/909,088B
/ CURRENT FILING DATE: 2001-07-18
/ PRIOR APPLICATION NUMBER: PCT/US00/04414
/ PRIOR FILING DATE: 2000-02-22
/ PRIOR APPLICATION NUMBER: US 60/143,048
/ PRIOR FILING DATE: 1999-07-07
/ PRIOR APPLICATION NUMBER: US 60/145,698
/ PRIOR FILING DATE: 1999-07-26
/ PRIOR APPLICATION NUMBER: US 60/146,222
/ PRIOR FILING DATE: 1999-07-28
/ PRIOR APPLICATION NUMBER: PCT/US99/20594
/ PRIOR FILING DATE: 1999-09-08
/ PRIOR APPLICATION NUMBER: PCT/US99/20944
/ PRIOR FILING DATE: 1999-09-13
/ PRIOR APPLICATION NUMBER: PCT/US99/21090
/ PRIOR FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: PCT/US99/21547
/ PRIOR FILING DATE: 1999-09-15
/ PRIOR APPLICATION NUMBER: PCT/US99/23089
/ PRIOR FILING DATE: 1999-10-05
/ PRIOR APPLICATION NUMBER: PCT/US99/28214
/ PRIOR FILING DATE: 1999-11-29
/ PRIOR APPLICATION NUMBER: PCT/US99/28313
/ PRIOR FILING DATE: 1999-11-30
/ PRIOR APPLICATION NUMBER: PCT/US99/28564
/ PRIOR FILING DATE: 1999-12-02
/ PRIOR APPLICATION NUMBER: PCT/US99/28565
/ PRIOR FILING DATE: 1999-12-02
/ PRIOR APPLICATION NUMBER: PCT/US99/30095
/ PRIOR FILING DATE: 1999-12-16
/ PRIOR APPLICATION NUMBER: PCT/US99/30911
/ PRIOR FILING DATE: 1999-12-20
/ PRIOR APPLICATION NUMBER: PCT/US99/30999
/ PRIOR FILING DATE: 1999-12-20
/ PRIOR APPLICATION NUMBER: PCT/US00/00219
/ PRIOR FILING DATE: 2000-01-05
/ NUMBER OF SEQ ID NOS: 423
/ SEQ ID NO 34
/ LENGTH: 915
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-909-088B-34

```

Query Match	12.0%	Score	350.5	DB	10	Length	915
Best Local Similarity	34.3%	Pred.	No. 8e-15				
Matches	85	Conservative	37	Mismatches	107	Indels	19
Gaps							14
QY	245	CSVENGCCCEHACNAIPCAPRCQCPAGAAALQADGRSCTA-SATQSCNDLCEHFCVNPDPQP	303				
Db	283	CAMEDHNCEOLCVNVPGSFVCQYSGYALAEADGKRCAVDYCASENHGCHECV-NAD--	339				
QY	304	GSYSCMCETCYRLAAADQHRCEDYDDCILEPSPQRCVNTQGGFECHCYPNYDL-VGEC	362				
Db	340	GSYLCOCHGEFALNPDEKTCIRNYCALNPGCBHECHVMNEESYICRHRGTYTLDPNGKT	399				
QY	363	VEPVDPCFRAN--CEYOCQPLN-QTSYLCVCAEGFADIPHEPHRCQM--FC--NQTACPA	415				
Db	400	CSRYDHCAQDHHGEQLC--LNTEDSFVCCSEGL-INEDLKTCRSVDYCLLSDHGCEY	456				
QY	416	DC-DPNTQASCEPEGYIL-DDGFCIDIDBCENGGF-CGCVCHNLPTGTEICGPPSAL	472				
Db	457	SCVNMDBRSFACQCPEGHVLRSDEKTKACKLDSALGDHGCHEHSCVSSDESFCVQCF-EGYI	515				
QY	473	VRHIGTDC	480				
Db	516	LREDGKTC	523				

```

RESULT 11
US-09-747-371-2
; Sequence 2, Application US/09747371
; Patent No. US20020006616A1
; GENERAL INFORMATION:
; APPLICANT: Gish, Kurt
; APPLICANT: Mack, David
; TITLE OF INVENTION: No. US20020006616a1el Methods of Diagnosing Breast Cancer, Com
; TITLE OF INVENTION: Screening for Breast Cancer Modulators
; FILE REFERENCE: A-69028/DJB/JJD
; CURRENT APPLICATION NUMBER: US/09/747,371
; CURRENT FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: PCT/ US/00/06952
; PRIOR FILING DATE: 2000-03-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 999
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-747-371-2

```

```

Query Match          11.4%; Score 333; DB 10; Length 999;
Best Local Similarity 25.5%; Pred. No. 1.le-13;
Matches 107; Conservative 34; Mismatches 125; Indels 154; Gaps 20;

QY 182 PGAAAVASITYGTPFAARGADFQALPVGSSAAVAPLGLQLMCTAPPGAVOGHWARE--- 238
      ||||| : : : | ||| : ||| : : |
Db 9 PGAAWVLLLLLLLP-----PL-LLAGAVPPGRGAAQGPQEDVD 47

QY 239 --APGAWDCSVENGCCBEHA---CNAIPGRPCOCAGAAQAADGRSCTASATQSCNDL-- 291
      ||||| : : : | ||| : ||| : : |
Db 48 ECAGGLDDC-----HADALCONTSTYKCSCKPG--YQGEGRQC-EDIDECGNELNG 96

QY 292 -CEHFCVPNPDPQGSYSCMCTGYRLAADQHRCEBDVDDCILEPSPQPCQVNTQGGFECH 350
      ||||| : : : | ||| : ||| : : |
Db 97 GCVHDCL--NIPGNRYCTCFDGFMLAHGHNCLDVDECLENNCGCOHTCVNVWGSYECC 153

QY 351 CYPNYDLVGE--CVEPVD---PCFRAN--CEYQCQPLNQTSYLCVCAEGF----- 394
      | : : : : | : : : : | : : : : |
Db 154 CKEGFFLSDNQHTCIHRSEGLSCMNKDHGCSHCICKEAPRGSVACRCRPFELAKNORDC 213

QY 395 -APIPHEPHRCQMECNOTA-----412
      | ||| : : ||
Db 214 ILTCNHGNGGCQHSDDTADGPECSPQYKMHDTGRSCLEREDTVLEVTESNTTSVVDG 273

QY 413 -----CPADC-DRNTQASCECPGYILD-DGFTCTDIDCE--NGG 449
      ||||| : : : | ||| : ||| : : |
Db 274 DKRVKRLLMETCAVNNGGCDRTCKDFTSTGVHSCPVGFTQLQDGGTKDIDECQTRNG 333

QY 450 FCSGVCHNLPGTFECIGP-----DSALVRHIGT---DCDSG 483
      ||||| : : ||
Db 334 -CDHFCKNIVGSFDGCKKGFLLTDEKSCQDVDECSLDRFCDHSCINHPGTFACACNRG 392

RESULT 12
US-09-898-570-16
; Sequence 16, Application US/09898570
; Patent No. US20020123612A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: 15966-776CIP
; CURRENT APPLICATION NUMBER: US/09/898,570
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/198,293
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 60/198,645

```



Query Match	Best Local Similarity	Score	DB	Length	Indels	Gaps
11.3%;	25.68;	328.5;	DB 10;	Length 1009;		
Matches 90;	Conservative 34;	Mismatches 126;	Indels 101;	Gaps 14;		
QY 235 WAREAPGAWDC--SVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSC--TASATQSCND 290						
DB 23 YSKAAQDVDECVGTGDNCHIDAICQNTPRSYKCIKSG--YTGDKGKCKDVDECEREDNA 80						
QY 291 LCEHFVCPNPDPGSGYSYSCMCETGYRLAADQHRCEVDVDCILEPSPCPQRCVNTQGGFECH 350						
DB 81 GCVHDCV---NIPGNRYCTCYDGFHLADHGNCLDVDECAEGNGCQCSVNMMSGYECH 137						
QY 351 CYPNYDLVDGE--CPEPVDP---CFRAN--CEYQCQPLNQTSYLCVCAEGFAPIPHEPHR 403						
DB 138 CREGFFLSDNQHTCIQRPESGMCMKNKHGCAHICRETIPKGGIACECRPGF-ELIFKNQRD 196						
QY 404 COMFC-----NIPENTRCTCYDGFHLADHGNCL 408						
DB 197 CKLTCNYNGGCGQHTCDDTEQGRGCGCHIKFVLTGDKTCIGERRLEQHIPTQAVSNETC 256						
QY 409 --NOTACPADC-DPNTQASCEPGEYILD-DGFICTDIDECB-NGGFCSGVCHNLPGTPE 463						
DB 257 AVNNGGCDKCHDAATGVHCTCPVGMFLQPRDKTKIDECRLNNGGCDHICRNTVGSFE 316						
QY 464 CIGCPDSALV-----RHIGTDCDSKGVGDGSDSGSGEPPTGGS 503						
DB 317 CSCKKGKLLINERNCQDIDERSFDTCDHICVN-----TPGS 354						
RESULT 13						
US-09-898-570-14						
Sequence 14, Application US/09898570						
Patent No. US20020123612A1						
GENERAL INFORMATION:						
APPLICANT: GERLACH, VALERIE L.						
APPLICANT: ELLERMAN, KAREN						
APPLICANT: MACDOUGALL, JOHN R.						
APPLICANT: SMITHSON, GLENDA						
TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF USING THE SAME						
FILE REFERENCE: 15966-776CIP						



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:19:38 ; Search time 34.5 seconds  
(without alignments)  
1992.966 Million cell updates/sec

Title: US-09-509-994-2

Perfect score: 2916  
Sequence: 1 MLGVILVGLALAGLGFAP.....PSPETGSLTPPAVLVHSG 516

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2916	100.0	516	20	AAV09348
2	2916	100.0	516	21	AAV83935
3	2912	99.9	516	13	AAE22016
4	2912	99.9	516	20	AAV09347
5	2912	99.9	516	21	AAV83934
6	2912	99.9	516	21	AAV69529
7	2912	99.9	516	12	AAV11534
8	2912	99.9	575	14	AAV41806
9	2912	99.9	575	14	AAV43031
10	2912	99.9	575	23	AAE23026

11	2912	99.9	575	23	AAE17521	Human full-length
12	2910	99.8	515	12	AAE10617	Soluble thrombomod
13	2910	99.8	575	9	AAV82070	Human thrombomodul
14	2908	99.7	516	13	AAE22018	Human thrombomodul
15	2908	99.7	575	14	AAE31572	Human thrombomodul
16	2904	99.6	516	13	AAE22017	Human thrombomodul
17	2903	99.6	575	13	AAE20639	Human urinary thro
18	2879	98.7	575	20	AAW73970	Human thrombomodul
19	2862	98.1	575	13	AAE22189	Sequence of thromb
20	2826	96.9	498	16	AAE84185	Human derived thro
21	2826	96.9	498	18	AAW01600	Thrombomodulin TME
22	2826	96.9	498	21	AAE7402	Novel sugar chain-
23	2824	96.8	497	17	AAE94607	Human recombinant
24	2823	96.8	557	23	AAE23028	Human thrombomodul
25	2823	96.8	557	23	AAE23031	Human thrombomodul
26	2823	96.8	557	23	AAE17525	Human thrombomodul
27	2823	96.8	557	23	AAE17528	Human thrombomodul
28	2822	96.8	557	23	AAE17523	Human thrombomodul
29	2821	96.7	557	23	AAE17596	Human thrombomodul
30	2820	96.7	497	19	AAE69520	rSTM protein SEQ I
31	2820	96.7	557	23	AAE17530	Human thrombomodul
32	2820	96.7	557	23	AAE17598	Human thrombomodul
33	2819	96.7	557	23	AAE23029	Human thrombomodul
34	2819	96.7	557	23	AAE17526	Human thrombomodul
35	2818	96.6	557	23	AAE23030	Human thrombomodul
36	2818	96.6	557	23	AAE17527	Human thrombomodul
37	2818	96.6	557	23	AAE17529	Human thrombomodul
38	2818	96.6	557	23	AAE17593	Human thrombomodul
39	2818	96.6	557	23	AAE17595	Human thrombomodul
40	2817	96.6	557	23	AAE17597	Human thrombomodul
41	2817	96.6	557	23	AAE17591	Human thrombomodul
42	2817	96.6	557	23	AAE17592	Human thrombomodul
43	2817	96.6	557	23	AAE17594	Human thrombomodul
44	2816	96.6	557	23	AAE17524	Human thrombomodul
45	2811	96.4	554	23	AAE23027	Human thrombomodul

ALIGNMENTS

RESULT 1  
AAV09348  
ID AAV09348 standard; Protein; 516 AA.  
XX AC AAV09348;  
XX DT 08-JUL-1999 (first entry)  
XX DE Human thrombomodulin SEQ ID NO:2.  
XX KW Human; thrombomodulin; aqueous parenteral solution; storage;  
KW distribution; acute coronary syndrome; thrombosis; embolism;  
KW diabetes.  
XX OS Homo sapiens.  
XX PN WO9918994-A1.  
XX PD 22-APR-1999.  
XX PF 13-OCT-1998; 98WO-JP04609.  
XX PR 11-NOV-1997; 97JP-0308523.  
XX PR 15-OCT-1997; 97JP-0281659.  
XX PA (ASAH ) ASAH KASEI KOGYO KK.  
XX PI Murata T, Shimizu H, Tsuruta K, Yokozawa A, Yui M;  
XX DR WPI: 1999-277444/23.  
XX DR N-PSDB; AAX55880.  
XX PT Stable aqueous parenteral thrombomodulin solution - comprising

PT buffer and surfactant, useful for treating acute coronary syndrome,  
 PT thrombosis, embolism, and diabetes  
 XX  
 PS  
 XX  
 XX Claim 7; Page 90-92; 97pp; Japanese.  
 CC The present invention describes a method for maintaining the quality of  
 CC an aqueous parenteral solution of thrombomodulin comprising buffer and  
 CC surfactant aseptically filled in a case or syringe. Maintaining the  
 CC quality of an aqueous, parenteral thrombomodulin solution is  
 CC characterised in that the solution: (a) comprises soluble thrombomodulin,  
 CC a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically  
 CC filled into: (i) a case or (ii) a syringe without any empty space; and  
 CC (c) is kept in liquid form in storage and distribution and not frozen or  
 CC freeze dried. Thrombomodulin is used to treat acute coronary syndrome  
 CC (such as myocardial infarction, unstable angina and coronary artery  
 CC blockage), thrombosis (e.g. cerebral, vascular and peripheral blood  
 CC vessel thrombosis), embolism, peripheral blood vessel disorders (e.g.  
 CC Raynauds disease), arteriosclerosis, vasculitis (e.g. due to systemic  
 CC lupus erythematosus or harrett's syndrome), multiple organ failure,  
 CC disseminated intravascular coagulation, transient ischaemia, diabetes,  
 CC liver veno-occlusive diseases and deep vein thrombosis. The composition  
 CC is stable for a long period of time and can be stored and distributed in  
 CC ready to use form avoiding the problems of dissolution and accuracy when  
 CC preparing on demand. The present sequence represents human  
 CC thrombomodulin.  
 XX  
 SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 20; Length 516;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-163;  
 Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVGLALAGLGFPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
 DB 1 MLGVLVGLALAGLGFPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
 QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFQVWTDGNNNTSYS 120  
 DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFQVWTDGNNNTSYS 120  
 QY 121 RWARDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEHFHFPATCRPLAV 180  
 DB 121 RWARDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEHFHFPATCRPLAV 180  
 QY 181 EPGAAAASVITYGTFFAARGADFQALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240  
 DB 181 EPGAAAASVITYGTFFAARGADFQALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240  
 QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
 DB 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
 QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360  
 DB 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360  
 QY 361 ECVEPVDFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
 DB 361 ECVEPVDFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
 QY 421 TQASCECPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDPSALARHIGTDC 480  
 DB 421 TQASCECPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDPSALARHIGTDC 480  
 QY 481 DSGKVDGSGSGGEPPTPGSTLTTPPAVGLVHSG 516  
 DB 481 DSGKVDGSGSGGEPPTPGSTLTTPPAVGLVHSG 516

RESULT 2  
 AAY83935  
 ID AAY83935 standard; Protein; 516 AA.  
 XX

AC AAY83935;  
 XX 28-JUL-2000 (first entry)  
 XX Human thrombomodulin TMD protein #2.  
 DE  
 XX Human; thrombomodulin; vasculitis; protein C; thrombin.  
 KW  
 XX Homo sapiens.  
 OS  
 XX JP2000053582-A.  
 PN  
 XX 22-FEB-2000.  
 PD  
 XX 06-AUG-1998; 98JP-0222688.  
 PF  
 XX 06-AUG-1998; 98JP-0222688.  
 PR  
 XX (ASAH ) ASAH KASEI KOGYO KK.  
 PA  
 XX WPI: 2000-353249/31.  
 DR  
 XX N-PSDB; AAL10028.  
 XX  
 XX Treating agent for vasculitis contains peptide which promotes  
 PT activation of protein C by thrombin -  
 PT  
 XX Claim 4; Page 13-14; 18pp; Japanese.  
 PS  
 XX This sequence represents a human thrombomodulin protein. The invention  
 CC relates to a treating agent for vasculitis containing a peptide which  
 CC promotes activation of protein C by thrombin.  
 XX  
 SQ Sequence 516 AA;

Query Match 100.0%; Score 2916; DB 21; Length 516;  
 Best Local Similarity 100.0%; Pred. No. 7.8e-163;  
 Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVGLALAGLGFPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
 DB 1 MLGVLVGLALAGLGFPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
 QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFQVWTDGNNNTSYS 120  
 DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGRLGFQVWTDGNNNTSYS 120  
 QY 121 RWARDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEHFHFPATCRPLAV 180  
 DB 121 RWARDLNGAPLCGLPCVAVSAEAATVPSEPIWEEQQCEVKADGFLCEHFHFPATCRPLAV 180  
 QY 181 EPGAAAASVITYGTFFAARGADFQALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240  
 DB 181 EPGAAAASVITYGTFFAARGADFQALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240  
 QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
 DB 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
 QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360  
 DB 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360  
 QY 361 ECVEPVDFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
 DB 361 ECVEPVDFCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
 QY 421 TQASCECPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDPSALARHIGTDC 480  
 DB 421 TQASCECPGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDPSALARHIGTDC 480  
 QY 481 DSGKVDGSGSGGEPPTPGSTLTTPPAVGLVHSG 516  
 DB 481 DSGKVDGSGSGGEPPTPGSTLTTPPAVGLVHSG 516

Db	301	DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG	360
QY	361	ECVEPVDPFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQPCNQTACPADCDPN	420
Db	361	ECVEPVDPFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQPCNQTACPADCDPN	420
QY	421	TQASCECEGYILDDGFICTDIDECENGFCSCGVCHNLPGTFECICGPDALARIHGTDC	480
Db	421	TQASCECEGYILDDGFICTDIDECENGFCSCGVCHNLPGTFECICGPDALVRIHGTDC	480
QY	481	DSGKVDGSGSGEPPSPPTPGSTLTTPPAVGLVHSG	516
Db	481	DSGKVDGSGSGEPPSPPTPGSTLTTPPAVGLVHSG	516
RESULT 4			
AY	09347	standard; Protein: 516 AA.	
AC	AY09347;		
XX	08-JUL-1999	(first entry)	
DT	08-JUL-1999	(first entry)	
DE	Human thrombomodulin SEQ ID NO:1.		
XX	Human; thrombomodulin; aqueous parenteral solution; storage;		
KW	distribution; acute coronary syndrome; thrombosis; embolism;		
KW	diabetes.		
XX	Homo sapiens.		
XX	WO9918994-A1.		
PN	22-APR-1999.		
XX	13-OCT-1998; 98WO-JP04609.		
PF	11-NOV-1997; 97JP-0308523.		
XX	15-OCT-1997; 97JP-0281659.		
PR	(ASAH ) ASahi Kasei Kogyo KK.		
XX	Murata T, Shimizu H, Tsuruta K, Yokozawa A, Yui M;		
PI	WPI; 1999-277444/23.		
XX	N-PSDB; AAX55879.		
DR	Stable aqueous parenteral thrombomodulin solution - comprising		
XX	buffer and surfactant, useful for treating acute coronary syndrome,		
PT	thrombosis, embolism, and diabetes		
PT	Claim 6; Page 87-89; 97pp; Japanese.		
XX	The present invention describes a method for maintaining the quality of		
PS	an aqueous parenteral solution of thrombomodulin comprising buffer and		
XX	surfactant aseptically filled in a case or syringe. Maintaining the		
CC	quality of an aqueous, parenteral thrombomodulin solution is		
CC	characterised in that the solution: (a) comprises soluble thrombomodulin,		
CC	a buffer to maintain the pH at 5-7 and a surfactant; (b) is aseptically		
CC	filled into: (i) a case or (ii) a syringe without any empty space; and		
CC	(c) is kept in liquid form in storage and distribution and not frozen or		
CC	freeze dried. Thrombomodulin is used to treat acute coronary syndrome		
CC	(such as myocardial infarction, unstable angina and coronary artery		
CC	blockage), thrombosis (e.g. cerebral, vascular and peripheral blood		
CC	vessel thrombosis), embolism, peripheral blood vessel disorders (e.g.		
CC	Raynaud's disease), arteriosclerosis, vasculitis (e.g. due to systemic		
CC	lupus erythematosus or Barrett's syndrome), multiple organ failure,		
CC	disseminated intravascular coagulation, transient ischaemia, diabetes,		
CC	liver veno-occlusive diseases and deep vein thrombosis. The composition		
CC	is stable for a long period of time and can be stored and distributed in		
CC	ready to use form avoiding the problems of dissolution and accuracy when		
CC	preparing on demand. The present sequence represents human		
CC	thrombomodulin.		

Db	301	DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG	360
QY	361	ECVEPVDPFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQPCNQTACPADCDPN	420
Db	361	ECVEPVDPFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQPCNQTACPADCDPN	420
QY	421	TQASCECEGYILDDGFICTDIDECENGFCSCGVCHNLPGTFECICGPDALARIHGTDC	480
Db	421	TQASCECEGYILDDGFICTDIDECENGFCSCGVCHNLPGTFECICGPDALVRIHGTDC	480
QY	481	DSGKVDGSGSGEPPSPPTPGSTLTTPPAVGLVHSG	516
Db	481	DSGKVDGSGSGEPPSPPTPGSTLTTPPAVGLVHSG	516
RESULT 3			
AY	09347	standard; Protein: 516 AA.	
AC	AY09347;		
XX	08-JUL-1999	(first entry)	
DT	08-JUL-1999	(first entry)	
DE	Human thrombomodulin encoded by plasmid pSV2TMD1.		
XX	Thrombin binding site; blood clotting; TMD1 deleter.		
KW	Homo sapiens.		
OS	EP474273-A.		
PN	11-MAR-1992.		
XX	05-AUG-1991; 91EP-0202009.		
PF	03-AUG-1990; 90JP-0204978.		
XX	(ASAH ) ASahi Kasei Kogyo.		
PA	Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;		
PI	WPI; 1992-081820/11.		
XX	New polypeptide inhibits blood coagulation and platelet		
PT	aggregation - promotes thrombin catalysed protein C activation		
PT	for treating myocardial infarction, thrombosis, embolism, etc.		
XX	Example 1; Page 17; 112pp; English.		
PS	Plasmid M13mp19TMD3 (see AAR22013) was used as template for		
XX	site-directed mutagenesis. A 177bp fragment was deleted using the		
CC	"TMD1 deleter" oligonucleotide to give plasmid M13TMD1 which encodes		
CC	the first 516 N-terminal amino acids of human thrombomodulin. Plasmid		
CC	M13TMD1 was completely digested with HindIII and BamHI and a TMD1		
CC	fragment of ca. 1700bp was isolated. The fragment was ligated to		
CC	HindIII- and BglIII-cut plasmid pSV2-dhfr to give the recombinant		
CC	plasmid pSV2TMD1. See AAR22014-R22022 and AAR25072.		
XX	Sequence 516 AA;		
SQ	Query Match 99.9%; Score 2912; DB 13; Length 516;		
	Best Local Similarity 99.8%; Pred. No. 1.3e-162;		
	Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY	1	MLGVLVLGALAGLGFPAEPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGLHM	60
Db	1	MLGVLVLGALAGLGFPAEPQPGSQCVHEHDCFALYPGPATFLNASQICDGLRGLHM	60
QY	61	TVRSSVAADVLSLLNGDGGVRRRLWTGLQPPCGDPRKRLGRLGQWVTGDNNTSYS	120
Db	61	TVRSSVAADVLSLLNGDGGVRRRLWTGLQPPCGDPRKRLGRLGQWVTGDNNTSYS	120
QY	121	RWRLDLNGAPLCGLPLCVAVSAEATVPSEPIWEQQCEVKADGFLCFEHPFATCRPLAY	180
Db	121	RWRLDLNGAPLCGLPLCVAVSAEATVPSEPIWEQQCEVKADGFLCFEHPFATCRPLAY	180
QY	181	EPGAAAASVITYTTPFAARGADQALPVGSSAAVAPLGLQLMCTAPPAGVGHWAREAP	240
Db	181	EPGAAAASVITYTTPFAARGADQALPVGSSAAVAPLGLQLMCTAPPAGVGHWAREAP	240
QY	241	GAWDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCSTASATQSCNDLCEHFCVNP	300
Db	241	GAWDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCSTASATQSCNDLCEHFCVNP	300
QY	301	DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG	360

```

XX SQ Sequence 516 AA;
Query Match 99.9%; Score 2912; DB 20; Length 516;
Best Local Similarity 99.8%; Pred. No. 1.3e-162;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGVLVLAGALAGLGGFPAPAEPPQGGSCQVEHDCFALYEGPATFLNASQICDGLRGHLM 60
DB 1 MLGVLVLAGALAGLGGFPAPAEPPQGGSCQVEHDCFALYEGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRGLRGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRGLRGFQWVTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGFLCVAVSAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAV 180
DB 121 RWARLDLNGAPLCGFLCVAVSAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAV 180
QY 181 EPGAAAASVITYGTPTFAARGADFQALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPTFAARGADFQALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCBHFVCVNP 300
DB 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCBHFVCVNP 300
QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQPCVNTQGGFECCHCYPNYDLVDG 360
DB 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQPCVNTQGGFECCHCYPNYDLVDG 360
QY 361 ECVEPVPDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420
DB 361 ECVEPVPDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420
QY 421 TQASCECEPGYILDDGFICTDIDECENGFCGSGVCHNLPCTFECICGPDALARRHIGTDC 480
DB 421 TQASCECEPGYILDDGFICTDIDECENGFCGSGVCHNLPCTFECICGPDALARRHIGTDC 480
QY 481 DSGKVDGSDSGSGPPSPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSGPPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 5
AAY83934
ID AAY83934 standard; Protein; 516 AA.
XX AC AAY83934;
XX DT 28-JUL-2000 (first entry)
XX DE Human thrombomodulin TMD protein.
XX KW Human; thrombomodulin; vasculitis; protein C; thrombin.
XX OS Homo sapiens.
XX PN JP2000053582-A.
XX PD 22-FEB-2000.
XX PF 06-AUG-1998; 98JP-0222688.
XX PR 06-AUG-1998; 98JP-0222688.
XX PA (ASAH ) ASAH KASEI KOGYO KK.
XX DR WPI; 2000-353249/31.
XX DR N-PSDB; AAA10027.
XX PT Treating agent for vasculitis contains peptide which promotes
activation of protein C by thrombin

```

```

XX PS Claim 4; Page 10-12; 18pp; Japanese.
XX CC This sequence represents a human thrombomodulin protein. The invention
CC relates to a treating agent for vasculitis containing a peptide which
CC promotes activation of protein C by thrombin.
XX SQ Sequence 516 AA;
Query Match 99.9%; Score 2912; DB 21; Length 516;
Best Local Similarity 99.8%; Pred. No. 1.3e-162;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGVLVLAGALAGLGGFPAPAEPPQGGSCQVEHDCFALYEGPATFLNASQICDGLRGHLM 60
DB 1 MLGVLVLAGALAGLGGFPAPAEPPQGGSCQVEHDCFALYEGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRGLRGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRGLRGFQWVTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGFLCVAVSAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAV 180
DB 121 RWARLDLNGAPLCGFLCVAVSAEATVPSEPIWEEQCEVKADGFLCEHFFPATCRPLAV 180
QY 181 EPGAAAASVITYGTPTFAARGADFQALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPTFAARGADFQALPVGSSAAVAPLGLQLMCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCBHFVCVNP 300
DB 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCBHFVCVNP 300
QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQPCVNTQGGFECCHCYPNYDLVDG 360
DB 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQPCVNTQGGFECCHCYPNYDLVDG 360
QY 361 ECVEPVPDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420
DB 361 ECVEPVPDPCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNTACPADCDPN 420
QY 421 TQASCECEPGYILDDGFICTDIDECENGFCGSGVCHNLPCTFECICGPDALARRHIGTDC 480
DB 421 TQASCECEPGYILDDGFICTDIDECENGFCGSGVCHNLPCTFECICGPDALARRHIGTDC 480
QY 481 DSGKVDGSDSGSGPPSPPTPGSTLTTPPAVGLVHSG 516
DB 481 DSGKVDGSDSGSGPPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 6
AAY69529
ID AAY69529 standard; Protein; 516 AA.
XX AC AAY69529;
XX DT 10-APR-2000 (first entry)
XX DE Human thrombomodulin variant, SEQ ID NO:3.
XX KW Thrombomodulin; TM; soluble; affinity purification; cation exchange;
XX OS Homo sapiens.
XX PN JP11341990-A.
XX PD 14-DEC-1999.
XX PF 23-MAR-1999; 99JP-0077518.
XX PR 30-MAR-1998; 98JP-0084389.
XX PT

```

PA (ASAH ) ASahi KASEI KOGYO KK.  
XX WPI; 2000-101696/09.  
DR N-PSDB; AA255965.  
XX Preparation of highly pure soluble thrombomodulin - used as an  
PT antithrombotic agent and a thrombolytic agent  
XX  
PS Claim 9; Page 30-32; 36pp; Japanese.  
XX  
XX The invention relates to a novel method for the preparation of highly  
CC pure soluble thrombomodulin (TM) containing substantially no serum-  
CC derived or antibody-derived substance. The method comprises isolating  
CC soluble TM from unpurified serum via affinity purification using an  
CC anti-TM antibody. The soluble TM is then treated with a cation  
CC exchanger at a specific conductivity of 25 to 34 ms/cm and a pH of 3 to  
CC 4, and the fraction containing the soluble TM is isolated. The soluble  
CC TM can be used as a blood anticoagulant agent and a thrombolytic agent.  
CC This sequence represents a human thrombomodulin variant, designated  
CC SEQ ID NO:3.  
XX  
XX Sequence 516 AA;  
SQ  
Query Match 99.9%; Score 2912; DB 21; Length 516;  
Best Local Similarity 99.8%; Pred. No. 1.3e-162;  
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MGVLVGLGALAGLGFPAEPQPGSGQCVHDCFALYPGPATFLNASQICDGLRGHLM 60  
DB 1 MGVLVGLGALAGLGFPAEPQPGSGQCVHDCFALYPGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPKRLGRLGFQVMTGDNNTSYS 120  
DB 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPKRLGRLGFQVMTGDNNTSYS 120  
QY 121 RWARLDNGAPLCGPLCAVSAEAATVPSEPIWEEOQCEVKADGFLCEHFFPATCRPLAV 180  
DB 121 RWARLDNGAPLCGPLCAVSAEAATVPSEPIWEEOQCEVKADGFLCEHFFPATCRPLAV 180  
QY 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
DB 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
QY 241 GAWDCSVENGCGEHCACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
DB 241 GAWDCSVENGCGEHCACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG 360  
DB 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG 360  
QY 361 ECVEPVDPCFRANCEYQCPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420  
DB 361 ECVEPVDPCFRANCEYQCPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420  
QY 421 TQASCEPEGYILDDGICHTDIDCENGGFCGVCNHLPGTFECICGPDLSALRHIGTDC 480  
DB 421 TQASCEPEGYILDDGICHTDIDCENGGFCGVCNHLPGTFECICGPDLSALRHIGTDC 480  
QY 481 DSGKVDGDSGSGEPSPPTPGSTLTTPPAGLVHSG 516  
DB 481 DSGKVDGDSGSGEPSPPTPGSTLTTPPAGLVHSG 516  
RESULT 7  
AAR11534  
ID AAR11534 standard; Protein; 575 AA.  
XX  
AC AAR11534;  
XX  
XX 17-JUN-1991 (first entry)  
XX  
DE Human thrombomodulin type II polypeptide.

XX Thrombomodulin; chondroitin; protein C; thrombin; thrombolysis;  
KW chondroitin sulphate; anticoagulant; myocardial infarction.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..18  
FT Peptide /label= signal peptide  
FT Peptide 19..365  
FT Peptide /label= opt. N-terminal sequence  
FT Peptide 366..480  
FT Peptide /label= human thrombomodulin active fragment  
FT Peptide 481..516  
FT Peptide /label= opt. C-terminal sequence  
XX WO9104276-A.  
XX 04-APR-1991.  
XX 25-SEP-1990; 90WO-JP01234.  
XX 25-SEP-1989; 89JP-0246270.  
XX (ASAH ) ASahi KASEI KOGYO.  
XX Yamamoto S, Gomi K, Ogawa K;  
XX WPI; 1991-117478/16.  
XX New human thrombomodulin polypeptide contg. chondroitin - has  
XX anticoagulant platelet agglutination and thrombolytic activity  
XX Disclosure; fig 1; 80pp; Japanese.  
XX This human thrombomodulin polypeptide comprises a 115 residue active  
XX fragment with opt. flanking sequences (N-terminal- 347 amino acids;  
XX C-terminal- 36 amino acids). Additionally it has an N-terminal signal  
XX sequence and an attached sugar chain, contg. chondroitin or chondroitin  
XX sulphate. It promotes the activation of protein C by thrombin and has  
XX good anticoagulant, platelet aggregation inhibition and thrombolytic  
XX activities. It can therefore be used for treating cardiovascular  
XX diseases, eg arteriosclerosis or myocardial infarction.  
XX  
XX Sequence 575 AA;  
SQ  
Query Match 99.9%; Score 2912; DB 12; Length 575;  
Best Local Similarity 99.8%; Pred. No. 1.5e-162;  
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MGVLVGLGALAGLGFPAEPQPGSGQCVHDCFALYPGPATFLNASQICDGLRGHLM 60  
DB 1 MGVLVGLGALAGLGFPAEPQPGSGQCVHDCFALYPGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPKRLGRLGFQVMTGDNNTSYS 120  
DB 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPKRLGRLGFQVMTGDNNTSYS 120  
QY 121 RWARLDNGAPLCGPLCAVSAEAATVPSEPIWEEOQCEVKADGFLCEHFFPATCRPLAV 180  
DB 121 RWARLDNGAPLCGPLCAVSAEAATVPSEPIWEEOQCEVKADGFLCEHFFPATCRPLAV 180  
QY 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
DB 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240  
QY 241 GAWDCSVENGCGEHCACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
DB 241 GAWDCSVENGCGEHCACNAIPGAPRCOCAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG 360  
DB 301 DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECHCYPNYDLVDG 360  
QY 361 ECVEPVDPCFRANCEYQCPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCDPN 420

|||||  
Db 361 ECVEPVDPCFRANCEYQCQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
Qy 421 TQASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALARIHIGTDC 480  
Db 421 TQASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALARIHIGTDC 480  
Qy 481 DSGKVDGDSGSGEPPTPGSTLTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPTPGSTLTPPAVGLVHSG 516  
RESULT 8  
AAR41806  
ID AAR41806 standard; peptide; 575 AA.  
AC AAR41806;  
XX  
DT 30-MAR-1994 (first entry)  
XX  
XX Thrombomodulin.  
XX  
KW Transformation; fungus; blood coagulation; prevention; platelet;  
KW aggregation; thrombolytic activity; thrombolysis; human thrombomodulin;  
KW site-directed mutagenesis; promotion; protein C; activation; thrombin.  
XX  
OS Homo sapiens.  
XX  
XX JP05213998-A.  
XX  
XX 24-AUG-1993.  
XX  
XX 02-AUG-1991; 91JP-0282369.  
XX  
XX 03-AUG-1990; 90JP-0204978.  
XX  
XX 30-JUL-1991; 91JP-0189984.  
XX  
XX (ASAH ) ASAH CHEM IND CO LTD.  
XX  
XX WPI; 1993-299652/38.  
XX  
XX Novel polypeptide obtd. by culturing transformed fungus - having  
XX blood coagulation preventing, platelet aggregation preventing and  
XX thrombolytic activities  
XX  
XX Disclosure; Fig 55; 65pp; Japanese.  
XX  
XX Novel polypeptides, obtd. by culturing transformed fungus, have  
XX blood coagulation preventing, platelet aggregation preventing  
XX and thrombolytic activities.  
XX  
XX In an example, plasmid M13mp19MD3 (constructed from pSV2TMJ2  
XX (ATCC 67238) contg. a human thrombomodulin sequence) was subjected  
XX to site directed mutagenesis to prepare pSV2TMD7. Plasmid pSV2TMD7  
XX was transfected to COS-1 cells. The activity of promoting protein C  
XX activation by thrombin of the peptide produced by the transformed  
XX COS-1 cell was measured. The amt. of the peptide was determined.  
XX  
SQ Sequence 575 AA;  
Query Match 99.9%; Score 2912; DB 14; Length 575;  
Best Local Similarity 99.8%; Pred. No 1.5e-162;  
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 MLGVLVLGALALAGLGPAPAEPPQPGSQVCQVHDFCALYPGPATFLNASQICDGLRGHLM 60  
Db 1 MLGVLVLGALALAGLGPAPAEPPQPGSQVCQVHDFCALYPGPATFLNASQICDGLRGHLM 60  
Qy 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGFLRGFWVGTGDNNTSYS 120  
Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQPLPGCGDKRLGFLRGFWVGTGDNNTSYS 120  
Qy 121 RWARDLNGAPLCGFLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAY 180  
|||||

Db 121 RWARDLNGAPLCGFLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFFHFPATCRPLAY 180  
Qy 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHAREAP 240  
Db 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPGAVQGHAREAP 240  
Qy 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAACQADGRSCTASATQSCNDLCEHFCVPNP 300  
Db 241 GAWDCSVENGGEHACNAIPGAPRCQCPAGAACQADGRSCTASATQSCNDLCEHFCVPNP 300  
Qy 301 DOPGSYSCMETGYRLAADQHRCEVDVDCILEPSPQRCQVNTQGGFECHCYPNDLVDG 360  
Db 301 DOPGSYSCMETGYRLAADQHRCEVDVDCILEPSPQRCQVNTQGGFECHCYPNDLVDG 360  
Qy 361 ECVEPVDPCFRANCEYQCQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
Db 361 ECVEPVDPCFRANCEYQCQLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
Qy 421 TQASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALARIHIGTDC 480  
Db 421 TQASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALARIHIGTDC 480  
Qy 481 DSGKVDGDSGSGEPPTPGSTLTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPTPGSTLTPPAVGLVHSG 516  
RESULT 9  
AAR43031  
ID AAR43031 standard; protein; 575 AA.  
XX  
AC AAR43031;  
XX  
DT 16-MAY-1994 (first entry)  
XX  
XX Human thrombomodulin.  
XX  
XX Anticoagulant; platelet aggregation inhibitor;  
XX protein C; activation; thrombin; thrombomodulin;  
XX coagulation disorder; thrombosis; myocardial infarction;  
XX embolism; telangiectasis; arteriosclerosis obliterans;  
XX disseminated intravascular coagulation; DIC; angina pectoris;  
XX gestosis; transient ischaemic attack.  
XX  
OS Homo sapiens.  
XX  
XX WO9322447-A.  
XX  
XX 11-NOV-1993.  
XX  
XX 30-APR-1993; 93WO-JP00578.  
XX  
XX 01-MAY-1992; 92JP-0112903.  
XX  
XX (ASAH ) ASAH CHEM IND CO LTD.  
XX (ASAH ) ASAH KASEI KOGYO KK.  
XX  
XX Kondo S, Toma K, Zushi M;  
XX  
XX WPI; 1993-368806/46.  
XX  
XX Peptide with anticoagulant and platelet aggregation inhibitor  
XX activity - which promotes protein C activation by thrombin and is  
XX useful in treating coagulation disorders e.g. thrombosis  
XX  
XX Disclosure; Fig 1; 84pp; Japanese.  
XX  
XX New peptides (see AAR50069) are inhibitors of the blood coagulation  
XX and platelet aggregation activities of thrombin and promote the  
XX protein-C activation effect of thrombin. They can be produced  
XX efficiently in pure form by culture of appropriate transformants,  
XX and are useful in treatment of circulatory disorders such as  
XX myocardial infarction, thrombosis, embolism, telangiectasis,





```

|||||
Db 421 TQASCEPEGILDDGFICTDIDECENGFCGCHNLPGTFECICGPDALSALVRHIGTDC 480
Qy 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516
Db 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 11
AAE17521
ID AAE17521 standard; Protein; 575 AA.
AC AAE17521;
XX
XX 22-APR-2002 (first entry)
XX
XX Human full-length native thrombomodulin (TM).
XX
XX Human; thrombomodulin; TM analogue; protein C; thrombotic disease;
XX antithrombotic therapy; thrombin-activatable fibrinolysis inhibitor;
XX TAFI; myocardial infarction; disseminated intravascular coagulation;
XX DIC; deep vein thrombosis; septic shock; pulmonary embolism; angina;
XX acute respiratory distress syndrome; cancer; toxemia; septicemia;
XX circulatory disorder; coronary embolism; pulmonary embolism; stroke;
XX systemic coagulation disorder; immunosuppressive; cerebroprotective;
XX pregnancy; anticoagulant; thrombolytic; venous occlusive condition;
XX antibacterial; cardiant.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Peptide 1..18
XX /label= Signal_peptide
XX Protein 19..575
XX /note= "Human mature thrombomodulin"
XX Domain 19..244
XX /label= N-terminal_domain
XX Domain 245..480
XX /note= "6 EGF-like domains"
XX Modified-site 481..515
XX /note= "O-linked glycosylation site"
XX Domain 516..539
XX /label= Transmembrane_domain
XX Domain 540..575
XX /label= Cytoplasmic_domain
XX
XX WO200198352-A2.
XX
XX 27-DEC-2001.
XX
XX 20-JUN-2001; 2001WO-US19590.
XX
XX 21-JUN-2000; 2000US-213678P.
XX 12-JUN-2001; 2001US-0880484.
XX
XX (SCHD ) SCHERING AG.
XX
XX Light D, Morser MJ, Nagashima M;
XX
XX WPI; 2002-154622/20.
XX N-PSDB; AAD28182.
XX
XX Thrombomodulin analog with greater ability to potentiate
XX thrombin-mediated activation of protein C and lesser ability to
XX potentiate thrombin-mediated activation of thrombin-activatable
XX fibrinolysis inhibitor for treating thrombotic disease -
XX
XX Claim 1; Fig 4; 53pp; English.
XX
XX The invention relates to the design, production and use of thrombomodulin
XX (TM) analogues that have the ability to enhance the thrombin-mediated
XX activation of protein C which have a significantly reduced ability to
XX promote activation of thrombin-activatable fibrinolysis inhibitor (TAFI).

```

```

CC TM analogue is useful for treating thrombotic disease in human, by
CC administering it or a polynucleotide encoding it to the patient and
CC expressing the analogue in vivo. Composition comprising TM analogue is
CC useful for treating a thrombotic disease or condition in mammals. TM
CC analogue is useful in antithrombotic therapy. TM analogue is also useful
CC for treating diseases in which thrombus formation plays a significant
CC etiological role including myocardial infarction, disseminated
CC intravascular coagulation (DIC), deep vein thrombosis, septic shock,
CC pulmonary embolism, acute respiratory distress syndrome, unstable angina
CC and other arterial or venous occlusive conditions. TM analogue is also
CC useful for treating various circulatory disorders including coronary or
CC pulmonary embolism, stroke and systemic coagulation disorders including
CC DIC which is often associated with septicemia, certain cancers and
CC toxemia of pregnancy. The present sequence is human full-length native
CC thrombomodulin.
XX
XX Sequence 575 AA;
XX
XX Query Match 99.9%; Score 2912; DB 23; Length 575;
XX Best Local Similarity 99.8%; Pred. No. 1.5e-162;
XX Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 MLGVVLGALALAGLGFAPAPQPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
XX Db 1 MLGVVLGALALAGLGFAPAPQPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60
XX
XX QY 61 TVRSSVAADVISILLNGDGGVGRRLWIGLQPPCGDPRKPLRGFQWVTDNNTSYS 120
XX Db 61 TVRSSVAADVISILLNGDGGVGRRLWIGLQPPCGDPRKPLRGFQWVTDNNTSYS 120
XX
XX QY 121 RWARLDNGAPLCGLPCVAVSAAEATVPSEPIWEEOQCEVKADGFLCEPHFATCRPLAV 180
XX Db 121 RWARLDNGAPLCGLPCVAVSAAEATVPSEPIWEEOQCEVKADGFLCEPHFATCRPLAV 180
XX
XX QY 181 EPGAAAASVITYGTPTFAARGADFQALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240
XX Db 181 EPGAAAASVITYGTPTFAARGADFQALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240
XX
XX QY 241 GAWDCSVENGCGCEHACNAIPGAPRCOCAGALQADGRSCTASATQSCNDLCEHFCVNP 300
XX Db 241 GAWDCSVENGCGCEHACNAIPGAPRCOCAGALQADGRSCTASATQSCNDLCEHFCVNP 300
XX
XX QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILFSPQPCQVCVNTQGGFECCHYPNYDLVDG 360
XX Db 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILFSPQPCQVCVNTQGGFECCHYPNYDLVDG 360
XX
XX QY 361 ECVEPVDPFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFQACPDADCPN 420
XX Db 361 ECVEPVDPFRANCEYQCQPLNOTSYLCVCAEGFAPIPHEPHRCQMFQACPDADCPN 420
XX
XX QY 421 TQASCECEGYILDDGFICTDIDECENGFCGCHNLPGTFECICGPDALSALVRHIGTDC 480
XX Db 421 TQASCECEGYILDDGFICTDIDECENGFCGCHNLPGTFECICGPDALSALVRHIGTDC 480
XX
XX QY 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516
XX Db 481 DSGKVDGSDSGSGEPSPPTPGSTLTTPPAVGLVHSG 516
XX
XX RESULT 12
XX AAR10617
XX ID AAR10617 standard; Protein; 515 AA.
XX
XX AC AAR10617;
XX
XX DT 17-APR-1991 (first entry)
XX
XX DE Soluble thrombomodulin deriv.
XX
XX KW Thrombosis; anticoagulant; Protein C.
XX OS Homo sapiens.
XX

```

FH Key Location/Qualifiers
FT Peptide 1..16
FT /label= signal peptide
FT /note= "for low mol. wt deriv.; optionally present"
FT Peptide 1..18
FT /label= signal peptide
FT /note= "for high mol wt. deriv; optionally present"
XX
XX EP412841-A.
XX
XX
XX 13-FEB-1991.
XX
XX 10-AUG-1990; 90EP-0308826.
XX
XX 05-FEB-1990; 90US-0474870.
XX 11-AUG-1989; 89US-0393617.
XX
XX (ELIL ) ELI LILLY & CO.
XX
XX Bang NU, Grinnell BW, Hoskins JA, Moore RE, Parkinson JF;
XX
XX WPI: 1991-045960/07.
XX N-PSDB; AAQ10435.
XX
XX New soluble thrombomodulin derivatives - produced by recombinant
XX DNA techniques for use as anticoagulants and in treatment of
XX thrombosis.
XX
XX Claim 2; Page 45; 81pp; English.
XX
XX The sequence is that of a sol. thrombomodulin deriv. comprising the
XX signal peptide (optional), N-terminal, epidermal growth factor
XX homology region and the Ser/Thr-rich region, but lacking the trans-
XX membrane and cytoplasmic domains. The sequence was deduced DNA
XX obtd. from clone GHTM3A isolated from a human chromosome 20 library,
XX subcloned into pUC19 to give pGHTM3A. A Pvu MI fragment of this
XX vector was ligated with a linker to construct pUC18TM, which was
XX treated with BsmI to delete about 500 bp from the 3' end of the
XX gene. ( The deletion occurs at the Ser/Thr-rich/transmembrane domain
XX junction, effectively cleaving the transmembrane and cytoplasmic
XX domains.) The plasmid was recircularised to give pUC18TMd for use
XX in prodn of the recombinant sol. deriv. The protein reacts with
XX thrombin to activate the Protein C anticoagulant pathway, inhibit
XX thrombin, and can be used to treat or prevent thrombotic disorders.
XX
XX Sequence 515 AA;
XX
XX Query Match 99.8%; Score 2910; DB 12; Length 515;
XX Best Local Similarity 100.0%; Pred. No. 1.7e-162;
XX Matches 515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MLGVLVLGALALAGLGFPAPEPQPGSGQCVHEHDCFALYGPATFLNASQICDGLRGHLM 60
DB 1 MLGVLVLGALALAGLGFPAPEPQPGSGQCVHEHDCFALYGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISILLNGDGGVGRRLRWIGLQPLPGCGDKPKRLGFLRGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISILLNGDGGVGRRLRWIGLQPLPGCGDKPKRLGFLRGFQWVTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGPLCAVAVSAEATVPSEPIWEEOQCEVKADGFLCEFFHPATCRPLAV 180
DB 121 RWARLDLNGAPLCGPLCAVAVSAEATVPSEPIWEEOQCEVKADGFLCEFFHPATCRPLAV 180
QY 181 EPGAAAAVSTYGTPTFAARGADQALPVGSSAAVAPLGLQLMCTAPPVAGVQGHWAREAP 240
DB 181 EPGAAAAVSTYGTPTFAARGADQALPVGSSAAVAPLGLQLMCTAPPVAGVQGHWAREAP 240
QY 241 GAWDCSVENGCCERHACNAIPGAPRCQCPAGAAQAQDGRSCTASATQSCNDLCEHFCVNP 300
DB 241 GAWDCSVENGCCERHACNAIPGAPRCQCPAGAAQAQDGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLYDG 360

Db 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLYDG 360
QY 361 ECVEPVDFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNQACPADCDPN 420
Db 361 ECVEPVDFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNQACPADCDPN 420
QY 421 TQASCECPEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPPSALARHIGTDC 480
Db 421 TQASCECPEGYILDDGFICTDIDECENGFCGSGVCHNLPGTFECICGPPSALARHIGTDC 480
QY 481 DSGKVDGDSGSGEPPPTPGSTLTTPPAVGLVHS 515
Db 481 DSGKVDGDSGSGEPPPTPGSTLTTPPAVGLVHS 515
RESULT 13
AAP82070
ID AAP82070 standard; protein; 575 AA.
XX
XX AAP82070;
XX
XX 19-OCT-1990 (first entry)
XX Human thrombomodulin encoded by plasmid p2.1.
XX
XX thrombomodulin activity; Protein C; anticoagulant;
KW epidermal Growth Factor (EGF) domains.
XX
XX synthetic.
XX
XX WO8809811-A.
XX
XX 15-DEC-1988.
XX
XX 09-JUN-1988; 88WO-DK00089.
XX
XX 12-JUN-1987; 87DK-0002990.
XX
XX (NOVO ) NOVO INDUSTRI A/S.
XX
XX Nexo BA, Esper B;
XX
XX WPI: 1988-368626/51.
XX N-PSDB; AAP82026.
XX
XX Recombinant protein having thrombomodulin activity -
XX used in the therapeutic control of coagulation and the treatment
XX and prevent of thrombotic episodes
XX
XX Disclosure; ; p; English.
XX
XX Protein is encoded by plasmid p2.1 derived from human cell
XX line A549 known to express about 10000 molecules of thrombomodulin
XX per cell. Plasmid p2.1 showed a strong hybridisation signal with a
XX 60-mer bovine thrombomodulin probe. Thrombomodulins are used to
XX potentiate a patient's anticoagulant capacity.
XX See also AAN82027 and AAN82037.
XX
XX Sequence 575 AA;
XX
XX Query Match 99.8%; Score 2910; DB 9; Length 575;
XX Best Local Similarity 99.8%; Pred. No. 1.9e-162;
XX Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLGVLVLGALALAGLGFPAPEPQPGSGQCVHEHDCFALYGPATFLNASQICDGLRGHLM 60
DB 1 MLGVLVLGALALAGLGFPAPEPQPGSGQCVHEHDCFALYGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISILLNGDGGVGRRLRWIGLQPLPGCGDKPKRLGFLRGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISILLNGDGGVGRRLRWIGLQPLPGCGDKPKRLGFLRGFQWVTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGPLCAVAVSAEATVPSEPIWEEOQCEVKADGFLCEFFHPATCRPLAV 180

```

|||||
Db 121 RWRDLNGLAPLCGLCVAVSAAEATVPSEPIWEEQCEVKADAFLECFHFPPATCRPLAV 180
QY 181 EPGAAAASVTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
Db 181 EPGAAAASVTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
QY 241 GAWDCSVENGCEHACNAIFGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFVCPNP 300
Db 241 GAWDCSVENGCEHACNAIFGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFVCPNP 300
QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360
Db 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360
QY 361 ECVPEVDFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNOTACPADCDPN 420
Db 361 ECVPEVDFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNOTACPADCDPN 420
QY 421 QAASCEPEGYILDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALARIHIGTDC 480
Db 421 QAASCEPEGYILDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALARIHIGTDC 480
QY 481 DSGKVDGSDSGSEPPSPPTPGSTLTTPPAVGLVHSG 516
Db 481 DSGKVDGSDSGSEPPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 14
AAR22018
ID AAR22018 standard; Protein; 516 AA.
XX AC AAR22018;
XX DT 03-JUL-1992 (first entry)
XX DE Human thrombomodulin (1-516) with Asp367 substituted by Glu.
XX KW Mutant; thrombin binding site; blood clotting; Tm2 mutator.
XX OS Homo sapiens.
XX PN EP474273-A.
XX PD 11-MAR-1992.
XX PF 05-AUG-1991; 91EP-0202009.
XX PR 03-AUG-1990; 90JP-0204978.
XX PA (ASAH ) ASahi Kasei Kogyo.
XX Zushi M, Gomi K, Yamamoto S, Suzuki K, Matsuda A;
XX WPI; 1992-081820/11.
XX New polypeptide inhibits blood coagulation and platelet
XX aggregation - promotes thrombin catalysed protein C activation
XX for treating myocardial infarction, thrombosis, embolism, etc.
XX Example 1; Page 18; 112pp; English.
XX Plasmid M13TMD1 (see AAR22016) encoding the first 516 N-terminal
XX amino acids of human thrombomodulin was used as a template for
XX site-directed mutagenesis using the "Tm2 mutator" to produce
XX plasmid M13TMD2. In the mutant thrombomodulin encoded by the
XX plasmid, the wild-type Asp residue at position 367 is substituted
XX by an Glu residue. The activity of this mutant, truncated
XX thrombomodulin (i.e. D123dlu) in the activation of protein C was
XX compared to that of similarly truncated thrombomodulin with Asp
XX at position 367 (i.e. D123Asp). The activity of D123Glu was as high
XX as 1.2 times that of D123Asp. See AAR22013-R22022 and AAR25074.

```

```

SQ Sequence 516 AA;
Query Match 99.7%; Score 2908; DB 13; Length 516;
Best Local Similarity 99.6%; Pred. No. 2.3e-162;
Matches 514; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 MLGVLVGLALAGLGPAPAEQPGGQCVHEHDCFALYFGPATFLNASICDGLRHLM 60
Db 1 MLGVLVGLALAGLGPAPAEQPGGQCVHEHDCFALYFGPATFLNASICDGLRHLM 60
QY 61 TVRSSVAADVLSILLNGDGGVRRRLWIGLQPPCCGDKPLRGFQWVTGDNNTSYS 120
Db 61 TVRSSVAADVLSILLNGDGGVRRRLWIGLQPPCCGDKPLRGFQWVTGDNNTSYS 120
QY 121 RWRDLNGLAPLCGLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPPATCRPLAV 180
Db 121 RWRDLNGLAPLCGLCVAVSAAEATVPSEPIWEEQCEVKADGFLCEHFPPATCRPLAV 180
QY 181 EPGAAAASVTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
Db 181 EPGAAAASVTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
QY 241 GAWDCSVENGCEHACNAIFGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFVCPNP 300
Db 241 GAWDCSVENGCEHACNAIFGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFVCPNP 300
QY 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360
Db 301 DQPGSYSCMCTGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360
QY 361 ECVPEVDFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNOTACPADCDPN 420
Db 361 ECVPEVDFCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFQNOTACPADCDPN 420
QY 421 QAASCEPEGYILDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALARIHIGTDC 480
Db 421 QAASCEPEGYILDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALARIHIGTDC 480
QY 481 DSGKVDGSDSGSEPPSPPTPGSTLTTPPAVGLVHSG 516
Db 481 DSGKVDGSDSGSEPPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 15
AAR31572
ID AAR31572 standard; Protein; 575 AA.
XX AC AAR31572;
XX DT 27-MAY-1993 (first entry)
XX DE Human thrombomodulin.
XX KW Mutagenesis; site specific; nucleic acid constructs;
XX restriction site; introduction; removal.
XX OS Synthetic.
XX Key Location/Qualifiers
FH Peptide 1..20
FT /note= "signal peptide"
FT Domain 245..280
FT /note= "EGF-1"
FT Domain 288..323
FT /note= "EGF-2"
FT Domain 329..361
FT /note= "EGF-3"
FT Domain 368..404
FT /note= "EGF-4"
FT Domain 408..439
FT /note= "EGF-5"
FT Domain 445..480
FT /note= "EGF-6"

```

Search completed: December 16, 2002, 17:24:25  
Job time : 35.5 secs

```
FT Domain 481..515
FT Region /note="O-linked glycosylation domain"
FT 516..538
FT /note="stop transfer sequence"
FT 539..575
FT /note="cytoplasmic domain"
XX
PN WO9301282-A.
XX
XX 21-JAN-1993.
XX
XX 01-JUL-1992; 92WO-US05573.
XX
XX 01-JUL-1991; 91US-0724237.
XX
XX (BERL-) BERLEX LAB INC.
XX
XX Andrews WH, Morser MJ, Vilander LR;
XX WPI; 1993-045488/05.
XX
XX Site-specific mutagenesis of nucleic acid constructs - using an
PT oligo:nucleotide which changes a nucleotide and introduces or
PT removes a restriction site
XX
XX Example; Fig 3; 87pp; English.
XX
XX The sequence is that of the native human thrombomodulin showing the
CC six EGF-like domains.
XX
XX Sequence 575 AA;

Query Match 99.7%; Score 2908; DB 14; Length 575;
Best Local Similarity 99.8%; Pred. No. 2.5e-162;
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGVLVIGALALAGLGPAPAEPOGGSCQVEHDCPALYFGPATFLNASQICDGLRGHLM 60
Db 1 MGVLVIGALALAGLGPAPAEPOGGSCQVEHDCSALYFGPATFLNASQICDGLRGHLM 60
Qy 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPPCGDKPLRGPLRGFWVTGDNNTSYS 120
Db 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPPCGDKPLRGPLRGFWVTGDNNTSYS 120
Qy 121 RWARLDLNGAPLCGLPCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAV 180
Db 121 RWARLDLNGAPLCGLPCVAVSAAEATVPSEPIWEEQCEVKADGFLCEFHFPATCRPLAV 180
Qy 181 EPGAAAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAGVGHWAREAP 240
Db 181 EPGAAAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAGVGHWAREAP 240
Qy 241 GAWDCSVENGCGEHCACNAIPCAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGCGEHCACNAIPCAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
Qy 301 DQPGSYSCMCTGYRLAADQHRCEDVDVDCILESPSPQRCVNTQGGFECHCYPNYDLVDG 360
Db 301 DQPGSYSCMCTGYRLAADQHRCEDVDVDCILESPSPQRCVNTQGGFECHCYPNYDLVDG 360
Qy 361 ECVPEVPDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNQACADCDPN 420
Db 361 ECVPEVPDPCFRANCEYQCQPLNQTSYLCVCAEGFAPIPHEPHRCQMFNCNQACADCDPN 420
Qy 421 TQASCECEGYILDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALARIHIGTDC 480
Db 421 TQASCECEGYILDGFICTDIDECENGFCGSGVCHNLPGTFECICGPDSSALARIHIGTDC 480
Qy 481 DSGKVDGSDSGSGEPSPPTFGSTLTTPPAVGLVHSG 516
Db 481 DSGKVDGSDSGSGEPSPPTFGSTLTTPPAVGLVHSG 516
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:22:58 ; Search time 16.5 Seconds  
(without alignments)  
3006.386 Million cell updates/sec

Title: US-09-509-994-2  
Perfect score: 2916  
Sequence: 1 MLGVLVGLALAGLGFAP.....PSPTGSLTPPAVGLVHSG 516

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_73.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2912	99.9	575	1 THHUB	thrombomodulin pre
2	1903	65.3	577	2 A60501	thrombomodulin pre
3	1095.5	37.6	356	2 A25918	thrombomodulin - b
4	414	14.2	1574	2 T13954	MEGF6 protein - ra
5	414	14.2	1620	2 T12783	hypothetical prote
6	374	12.8	2907	2 A57278	fibrillin-2 precu
7	368	12.6	2871	2 A55367	fibrillin-1 - bovi
8	367.5	12.6	1184	2 A55184	fibulin-2 precu
9	366	12.6	2918	2 A54105	fibillin-2 precu
10	364	12.5	3002	2 A47221	fibillin-1 precu
11	362	12.4	2871	2 A55624	fibillin-1 precu
12	357	12.2	741	2 T46488	hypothetical prote
13	357	12.2	1221	2 A49457	fibulin-2 precu
14	332.5	11.4	1964	2 T09059	notch4 - mouse
15	331	11.4	589	2 T43210	fibulin-1D precu
16	330.5	11.3	1712	2 A38261	masking protein pr
17	324.5	11.1	689	2 T42760	fibulin, splice fo
18	324.5	11.1	712	2 T42990	fibulin, splice fo
19	322.5	11.1	1394	2 A35626	transforming growt
20	321	11.0	2321	2 S78549	notch3 protein - h
21	314.5	10.8	3507	2 T34513	hypothetical prote
22	313	10.7	798	2 T22793	hypothetical prote
23	312.5	10.7	1820	2 A55494	latent transformin
24	311	10.7	601	2 B36346	fibulin 1 precu
25	311	10.7	683	2 C36346	fibulin 1 precu
26	309	10.6	685	2 S78040	fibulin, splice fo
27	308	10.6	705	2 S34968	fibulin, splice fo
28	303.5	10.4	1251	2 A57293	latent transformin
29	297	10.2	2531	2 A46019	Notch-1 protein -

RESULT 1

THHUB

thrombomodulin precursor [validated] - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1988 #sequence\_revision 12-May-1995 #text\_change 15-Sep-2000  
C:Accession: A41442; A28307; A29680; A27073; JX0264; S3954

R:Shirai, T.; Shiojiri, S.; Ito, H.; Yamamoto, S.; Kusumoto, H.; Deyashiki, Y.; Maruy  
J. Biochem. 103, 281-285, 1988

A:Title: Gene structure of human thrombomodulin, a cofactor for thrombin-catalyzed ac

A:Reference number: A41442; MUID:88227901; PMID:2836377

A:Accession: A41442

A:Molecule type: DNA

A:Residues: 1-575 <SHI>

A:Cross-references: DDBJ:D00210; NID:g220126; PIDN:BAA00149.1; PID:g220127

R:Jackman, R.W.; Beeler, D.L.; Fritze, L.; Soff, G.; Rosenberg, R.D.

Proc. Natl. Acad. Sci. U.S.A. 84, 6425-6429, 1987

A:Title: Human thrombomodulin gene is intron depleted: nucleic acid sequences of the

A:Reference number: A28307; MUID:87317665; PMID:2819876

A:Accession: A28307

A:Molecule type: DNA; mRNA

A:Residues: 1-472, 'A', 474-575 <JAC>

A:Cross-references: GB:J02973; NID:g339658; PIDN:AAA61175.1; PID:g339659

R:Suzuki, K.; Kusumoto, H.; Deyashiki, Y.; Nishioaka, J.; Maruyama, I.; Zushi, M.; Kaw

EMBO J. 6, 1891-1897, 1987

A:Title: Structure and expression of human thrombomodulin, a thrombin receptor on end

A:Reference number: A29680; MUID:88004395; PMID:2820710

A:Accession: A29680

A:Molecule type: mRNA

A:Residues: 1-575 <SU2>

A:Cross-references: GB:X05495; NID:g37123; PIDN:CAA29045.1; PID:g736251

A:Experimental source: lung endothelium

A:Note: Part of this sequence, including the amino end of the mature protein, were de

R:Wen, D.; Dittman, W.A.; Ye, R.D.; Deaven, L.L.; Majerus, P.W.; Sadler, J.E.

Biochemistry 26, 4350-4357, 1987

A:Title: Human thrombomodulin: complete cDNA sequence and chromosome localization of

A:Reference number: A27073; MUID:88024950; PMID:2822087

A:Accession: A27073

A:Molecule type: mRNA

A:Residues: 1-472, 'A', 474-575 <WEN>

A:Cross-references: GB:M16552; NID:g339656; PIDN:AAB59508.1; PID:g339657

A:Experimental source: placenta

A:Note: Parts of this sequence were determined by protein sequencing

R:Yamamoto, S.; Mizoguchi, T.; Tamaki, T.; Ohkuchi, M.; Kimura, S.; Aoki, N.

J. Biochem. 113, 433-440, 1993

A:Title: Urinary thrombomodulin, its isolation and characterization.

A:Reference number: JX0264; MUID:93293792; PMID:8390446

A:Accession: JX0264

A:Molecule type: protein; mRNA

A:Residues: 19-472, 'A', 474-486 <YAM>

A:Experimental source: urine

A:Note: the urinary form appears to be identical with that circulating in plasma

R:Gerlitz, B.; Hassell, T.; Vlahos, C.J.; Parkinson, J.F.; Bang, N.U.; Grinnell, B.W.

Biochem. J. 295, 131-140, 1993  
 A:Title: Identification of the predominant glycosaminoglycan-attachment site in soluble serine.  
 A:Reference number: S38954; MUID:94029900; PMID:8216207  
 A:Accession: S38954  
 A:Molecule type: protein  
 A:Residues: 475-491, 'X', 493-494 <GER>  
 A:Note: the residue designated 'X' was determined to be a Ser with covalently bound chondroitin-6-sulfate.  
 R:Maininger, D.P.; Komives, E.A.  
 submitted to the Brookhaven Protein Data Bank, September 1995  
 A:Reference number: A67369; PDB:1Z4Q  
 A:Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR, residue  
 R:Tulinsky, A.; Mathews, I.I.  
 submitted to the Brookhaven Protein Data Bank, August 1994  
 A:Reference number: A52804; PDB:1H1T  
 A:Contents: annotation; X-ray crystallography, 3.0 angstroms, residues 426-442  
 R:Hrabal, R.; Komives, E.A.; Ni, F.  
 submitted to the Brookhaven Protein Data Bank, November 1995  
 A:Reference number: A63583; PDB:1FGD  
 A:Contents: annotation; conformation by (1)H-NMR, residues 427-444  
 R:Hrabal, R.; Komives, E.A.; Ni, F.  
 Protein Sci. 5, 195-203, 1996  
 A:Title: Structural resiliency of an EGF-like subdomain bound to its target protein, the EGF receptor.  
 A:Reference number: A58595; MUID:96276211; PMID:8745396  
 A:Contents: annotation; conformation by (1)H-NMR  
 C:Genetics:  
 A:Gene: GDB:THBD  
 A:Cross-references: GDB:119613; OMIM:188040  
 A:Map position: 20p11.2-20p11.2  
 A:Introns: #status absent  
 C:Complex: homodimer, urinary form  
 C:Function:  
 A:Description: Inhibits thrombin activation of fibrinogen; cofactor for thrombin activation.  
 A:Pathway: blood coagulation/moderation  
 A:Note: the membrane-bound form is located on the endothelium luminal surface of arteries.  
 A:Note: thrombin complexed with the membrane-bound form is subject to endocytosis.  
 C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology  
 C:Keywords: anticoagulant; beta-hydroxyasparagine; beta-hydroxyaspartic acid; blood coagulation protein  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-575/Product: thrombomodulin, membrane-bound form #status predicted <MAT>  
 F:19-575/Domain: extracellular #status predicted <EXT>  
 F:19-486/Product: thrombomodulin, urinary form #status experimental <MAU>  
 F:24-167/Domain: C-type lectin homology <LCH>  
 F:177-199/Region: PEST sequence  
 F:201-233/Region: PEST sequence  
 F:245-280/Domain: EGF homology <EG1>  
 F:288-323/Domain: EGF homology <EG2>  
 F:329-362/Domain: EGF homology <EG3>  
 F:369-404/Domain: EGF homology <EG4>  
 F:408-439/Domain: EGF homology <EG5>  
 F:445-480/Domain: EGF homology <EG6>  
 F:485-513/Region: PEST sequence  
 F:517-539/Domain: transmembrane #status predicted <TMN>  
 F:540-575/Domain: intracellular #status predicted <INT>  
 F:47,115,116,382,409/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:174,225,411,504/Binding site: carbohydrate (Thr) (covalent) #status predicted  
 F:245-256,252-265,267-280,288-296,292-308,310-323,329-340,336-349,351-362,369-378,374-388,398/Binding site: carbohydrate (Ser) (covalent) #status predicted  
 F:334,498/Binding site: erythro-beta-hydroxyasparagine (Asn) #status experimental  
 F:342/Modified site: chondroitin sulfate (Ser) (covalent) (partial) #status experimental  
 F:490,492/Binding site: chondroitin sulfate (Ser) (covalent) (partial) #status experimental

Query Match 99.9%; Score 2912; DB 1; Length 575;  
 Best Local Similarity 99.8%; Pred. No. 2.8e-166;  
 Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGVLVIGALAGLGPAPAEQPGSGQVHDCFCALYPGATFNLASQICDGLRGHLM 60  
 Db 1 MLGVLVIGALAGLGPAPAEQPGSGQVHDCFCALYPGATFNLASQICDGLRGHLM 60

QY 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPFGGDPKRLGFLRGFWVTGDNNTSYS 120  
 Db 61 TVRSSVAADVISLLNGDGGVRRRLWIGLQLPFGGDPKRLGFLRGFWVTGDNNTSYS 120

QY 121 RWARDLNGAPLCPLCVAVSAAEATVSPIMEEOCEVKADGFLCEHFHPPATCRPLAV 180  
 Db 121 RWARDLNGAPLCPLCVAVSAAEATVSPIMEEOCEVKADGFLCEHFHPPATCRPLAV 180  
 QY 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVAVQGHAREAP 240  
 Db 181 EPGAAAASVITYGTTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVAVQGHAREAP 240  
 QY 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATSCNDLCSEHFCVNP 300  
 Db 241 GAWDCSVENGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATSCNDLCSEHFCVNP 300  
 QY 301 DQGSYSQMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYNYDLVDG 360  
 Db 301 DQGSYSQMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECHCYNYDLVDG 360  
 QY 361 ECVEPVPDPCFRANCEYOCQPLNQTSYLCVCAEGFAPIPHEPHRCMFCNQTACPADCPN 420  
 Db 361 ECVEPVPDPCFRANCEYOCQPLNQTSYLCVCAEGFAPIPHEPHRCMFCNQTACPADCPN 420  
 QY 421 TQASCECEGYLDDGFICTIDIDECENGSGVCHNLPGTFECICGSDSALAHRTGDC 480  
 Db 421 TQASCECEGYLDDGFICTIDIDECENGSGVCHNLPGTFECICGSDSALAHRTGDC 480  
 QY 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516  
 Db 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516

RESULT 2  
 A60501  
 thrombomodulin precursor - mouse  
 N:Alternate names: fetomodulin  
 C:Species: Mus musculus (house mouse)  
 C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 16-Jul-1999  
 R:Accession: S08488; A32001; A60501  
 R:Dittman, W.A.; Majerus, P.W.  
 Nucleic Acids Res. 17, 802, 1989  
 A:Title: Sequence of a cDNA for mouse thrombomodulin and comparison of the predicted  
 A:Reference number: S08488; MUID:89128454; PMID:2536925  
 A:Accession: S08488  
 A:Molecule type: mRNA  
 A:Residues: 1-577 <DIT>  
 A:Cross-references: EMBL:X14432; NID:g54781; PID:CAA32597.1; PID:g54782  
 R:Dittman, W.A.; Kumada, T.; Sadler, J.E.; Majerus, P.W.  
 J. Biol. Chem. 263, 15815-15822, 1988  
 A:Title: The structure and function of mouse thrombomodulin. Phorbol myristate acetat  
 A:Reference number: A32001; MUID:89008498; PMID:2844823  
 A:Accession: A32001  
 A:Molecule type: mRNA  
 A:Residues: 97-577 <DI2>  
 A:Cross-references: GB:J04060  
 R:Imada, S.; Yamauchi, H.; Nagumo, M.; Katayanagi, S.; Iwasaki, H.; Imada, M.  
 Dev. Biol. 140, 113-122, 1990  
 A:Title: Identification of fetomodulin, a surface marker protein of fetal development  
 A:Reference number: A60501; MUID:90292331; PMID:2162790  
 A:Accession: A60501  
 A:Molecule type: protein  
 C:Residues: 19-22; 330-343; 479-489; 545-555; 562-575 <IMA>  
 C:Comment: Thrombomodulin binds to and internalizes with thrombin. It is also a cofac  
 C:Superfamily: thrombomodulin; C-type lectin homology; EGF homology  
 C:Keywords: anticoagulant; membrane protein; phosphoprotein; receptor  
 F:24-165/Domain: C-type lectin homology <LCH>  
 F:244-279/Domain: EGF homology <EG1>  
 F:287-322/Domain: EGF homology <EG2>  
 F:328-361/Domain: EGF homology <EG3>  
 F:368-403/Domain: EGF homology <EG4>  
 F:407-438/Domain: EGF homology <EG5>  
 F:444-479/Domain: EGF homology <EG6>

Query Match 65.3%; Score 1903; DB 2; Length 577;  
 Best Local Similarity 65.6%; Pred. No. 3.1e-106;





A:Molecule type: DNA  
A:Residues: 1-1620 <WIL>  
A:Cross-references: EMBL:AL110498; NID: e1542303; PIDN: CAB54471.1; CESP: Y64G10A.f  
A:Experimental source: clone Y64G10A  
C:Genetics:  
A:Gene: CESP: Y64G10A.f  
A:Introns: 77/1; 116/1; 198/1; 282/1; 365/1; 425/1; 548/1; 559/1; 601/1; 625/1;  
Query Match 14.2%; Score 414; DB 2; Length 1620;  
Best Local Similarity 27.0%; Pred. No. 2.7e-17;  
Matches 128; Conservative 38; Mismatches 176; Indels 132; Gaps 20;  
QY 119 YSRWRLDLNGLPGLCPCLVAVSAE-----ATVSEPTWEEOQCE--VKADGFICEF 169  
DB 56 YLRFARFRSGGCKCLLRVQANCSADLCHNGGTCVPSFHNNDNQVCEPCVFTGACQY 115  
QY 170 HPPATCRPLAVEPGAAAAAVSTYGTTPAARGADFOALPGVSSAAVAPLGLQLMCTAPPG 229  
DB 116 D-ANEC--MANNNGGCEHCVN-TIGTYI-----CRWPG 145  
QY 230 AVOGHWAERAPGAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAALQADGRSCTASATQSC- 288  
DB 146 FELSGDNGTCSDDIDCAVNSGCDRCVNSPGGFRCDPCDLYLHADGRTG--GKVTSCS 203  
QY 289 --NDLCEHFCVPPNPQPGSYSCMCETGYRLAADOHQRCEDVDDCILEPSPCQPCVNTQGG 346  
DB 204 TDNGGCEHC--ENDSNGEFYRCRVGPKLSENKRSQVDPVPCFDNKGCGCOHHCNNHR 262  
QY 347 FECHYCPYNDL-VDCCEVEPVDPFCFRAN-CBYCQCPNLSYLCVCAEGFAPIPHEPHRC 404  
DB 263 AQCCQYGFELHLSYDRSCVDDECAKNGGCEHCENKGT-YRCKREGY-QLGRDGRTC 320  
QY 405 QMF-----CNQACPADC--DPTQASCEPCEGYL----- 433  
DB 321 EEMLGCGVGGGCGHQDCYDQDGGHVCKCRNGYILANDQKLCHDNTSTVIHARAPLWD 380  
QY 434 -----DDGF-----ICDIDDEC-ENGFGCSGVC 455  
DB 381 SYETVTCVPTDITLTKLCHMLDLSGHVQCFDGDGYELIDSKFCQDINFECHENNNGDSQIC 440  
QY 456 HNLPGTFECICGPDSSALAHICT-----DCDS-----GKVDGDSGSGEP 495  
DB 441 VNLGASVECCQCKPFLMKRKTCTEDISECSNNGGCGEQICSNQEGVMCSCEP 494  
RESULT 6  
A57278  
fibulin-2 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 23-Feb-1996 #sequence\_revision 23-Feb-1996 #text\_change 02-Aug-2002  
C:Accession: A57278  
R:Zhang, H.; Hu, W.; Ramirez, F.  
J. Cell Biol. 129, 1165-1176, 1995  
A:Title: Developmental expression of fibulin genes suggests heterogeneity of extracellular matrix  
A:Reference number: A57278; MUID: 95263670; PMID: 7744963  
A:Accession: A57278  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2907 <ZHA>  
A:Cross-references: NID: g762830; PIDN: AAA74908.1; PID: g762831  
C:Superfamily: fibulin 1; EGF homology  
F: 1239-1274/Domain: EGF homology <EGF1>  
F: 2488-2523/Domain: EGF homology <EGF>  
Query Match 12.8%; Score 374; DB 2; Length 2907;  
Best Local Similarity 32.9%; Pred. No. 1.1e-14;  
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;  
QY 244 DCSVNGGCEHCACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNLCHEFCVNPDD-- 301  
DB 1238 ECMIMNGCDTQCTNSGSEYECSCSEGYALMPDGRSCA-----DIDE--CENNPDI 1287  
QY 302 -----QPGSYSCMCETGYRLAADOHQRCEDVDDCILEPSPCQPCVNTQGGFECHCYP 353

DB 1288 DGQCQTNIPGEYRCLCYDGFMA5MDMKTCLDVNECDLNICMFGECENTKGSFICHCOL 1347  
QY 354 NYDLVDGE--CVBPVDPG--FRANCEYQCOPLN-QTSYLCVCAEGFA-----PIP 398  
DB 1348 GYSVKGGTGTCTD-VDECEIGHNCDMHASCLNVPGSFKSCREGWVGNGIKCIDLDECA 1406  
QY 399 HEHRCQMFNCQACPADCDPNTQAS--CECPGYLLDDGFTCTDIDE-----CENG- 448  
DB 1407 NGTHQCSI-----NAQC-VNTPGSRACSEGF-TGDFGFTCSVDDECAENTLNCENGQ 1457  
QY 449 -----GF-----CS-----GVCHNLPCTFECICGPDSSALA 473  
DB 1458 CLNVPAGYRCECEMGFTPASDSRSCQDIDECSPQNCVFTGTCNNLPGMFHCICDDGYELD 1517  
QY 474 RHIG--TDICD 481  
DB 1518 RTGGNCTDID 1527  
RESULT 7  
A55567  
fibulin 1 - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 02-Aug-2002  
C:Accession: A55567  
R:Wijstra, D.J.; Li, L.; Potter, K.A.; Womack, J.; Byers, P.H.  
Genomics 23, 480-485, 1994  
A:Title: Sequence of the coding region of the bovine fibulin cDNA and localization  
A:Reference number: A55567; MUID: 95137597; PMID: 7835900  
A:Accession: A55567  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-2871 <TIL>  
A:Cross-references: GB: L28748; NID: g508427; PIDN: AAA74122.1; PID: g508428  
C:Superfamily: fibulin 1; EGF homology  
F: 1201-1236/Domain: EGF homology <EGF>  
Query Match 12.6%; Score 368; DB 2; Length 2871;  
Best Local Similarity 31.5%; Pred. No. 2.4e-14;  
Matches 111; Conservative 26; Mismatches 109; Indels 106; Gaps 20;  
QY 244 DCSVNGGCEHCACNAIPGAPRCQCPAGAAALQADGRSCTASATQSCNLCHEFCVNP-- 300  
DB 1200 ECSIMNGCCTCTNSGSEYECSCQPGFALMPDQRSC-----DIDE--CEDNPNTC 1249  
QY 301 -----DQPGSYSCMCETGYRLAADOHQRCEDVDDCILEPSPCQPCVNTQGGFECHCYP 353  
DB 1250 DGQCQTNIPGEYRCLCYDGFMA5MDMKTCLDVNECDLNICLSGTCENTKGSFICHCDM 1309  
QY 354 NYDLVDGE--CVBPVDPG--FRANCEYQCOPLNQT-SYLCVCAEGFA-----PIP 398  
DB 1310 GYSKGGKGTCTD-INECEIGHNCDRHAVCTNTAGSKSCSPGWIGDKICTDLDECS 1368  
QY 399 HEHRCQMFNCQACPADCDPNTQAS--CECPGYLLDDGFTCTDIDE--EN----- 447  
DB 1369 NGTHMCSQH-----ADC-KNTMGSYRCLCKEGY-TGDFGFTCTDIDECSENLNGNGQ 1419  
QY 448 -----GCF-----CS-----GVCHNLPCTFECICGPDSSALA 473  
DB 1420 CLNAPGGYRCECEMGFTPASDSRSCQDIDECSPQNCVFTGTCNNLPGLFRCCEIGEYELD 1479  
QY 474 RHIG-----TDCDSGK-VDDGDSGSGEPSPPTPGSTLTTPPAVGLV 513  
DB 1480 RSGNGCTDNECLDPTTCISGNCVNTPGSVTCDPCPD-----FELNPTRVGCV 1527  
RESULT 8  
A55184  
fibulin-2 precursor - human  
N: Alternate names: protein DKFzps586A1519.1  
C: Species: Homo sapiens (man)  
C: Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 02-Aug-2002

C:Accession: A55184; T08744  
R:Zhang, R.Z.; Pan, T.C.; Zhang, Z.Y.; Mattei, M.G.; Timpl, R.; Chu, M.L.  
Genomics 22, 425-430, 1994  
A:Title: Fibulin-2 (FBLN2): human cDNA sequence, mRNA expression, and mapping of the gene  
A:Reference number: A55184; MUID:95104855; PMID:7806230  
A:Accession: A55184  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1184 <ZHA>  
A:Cross-references: GB:X82494; NID:9575232; PIDN:CAA57876.1; PID:g575233  
R:Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, March 1999  
A:Reference number: Z16471  
A:Accession: T08744  
A:Molecule type: mRNA  
A:Residues: 656-719, 'ODECLMGADHCSRRCVNTLGSFYCVNHTVLCADGYILNAHRKCYD', 720-853, 'T', 855-1184  
A:Cross-references: EMBL:AL050095  
A:Experimental source: adult uterus; clone DKF2p586A1519  
C:Genetics:  
A:Gene: GDB:FBLN2  
A:Cross-references: GDB:293037; OMIM:135821  
A:Map position: 3p25-3p24  
A:Note: DKF2p586A1519.1  
C:Superfamily: fibulin-2; EGF homology  
C:Keywords: alternative splicing; extracellular matrix  
F:1-27/Domain: signal sequence #status predicted <SIG>  
F:28-1184/Product: fibulin-2 protein #status predicted <MAT>  
F:905-941/Domain: EGF homology <EGF>  
Query Match 12.6%; Score 367.5; DB 2; Length 1184;  
Best Local Similarity 30.0%; Pred. No. 1.2e-14;  
Matches 119; Conservative 29; Mismatches 124; Indels 125; Gaps 21;  
QY 154 EQQCEVKAD-GFLCEPHFATC-----RPLAV-----EPGAAAAVSITYGTPFAARG 201  
Db 531 EGQSCSNPLNLGYPN-HVMSLSCGEEPLIVPEVRPPPEPAAAPRRVS-----EAE 582  
QY 202 ADFQALPVGSSAAV---APLGLQLMCTAPPGAVQGHWAREAPGAWDCSVENGCGEACNA 258  
Db 583 AGRALSIGTAEPLNSLPGDDQDECLLLPGEL-----COHLGIN 622  
QY 259 IPGAPRCQCPAGAAQLADGRSC-----TASAT-----QS 287  
Db 623 TVGSYHCACFPGLSDQDGRTPCRPEGHPPQPEAPQEPALKSEFSQVASTIPTPLPQPN 682  
QY 288 CND--LCBHFVCPNDPGVSCMCETGYRLAADQHRCEVDVDCILEPSPCP--QRCVNT 343  
Db 683 CKDNGPKQVC---STVGSAICSPFGYATMADGVSCEDINECVTDLHTCSRGEHCYNT 739  
QY 344 QGGFECH---CYPNYDLVDGCEVPEVDPFCFRANCEYOCQPLNQTSYLCV----- 389  
Db 740 LGSFHCYKALTCPEGYALKDGE-EDVDEC--AMGTHTCOP-----GFLCNTKGSFYCOA 792  
QY 390 ---CAEGFAPTPH-----EPHRCOMFCNOTACPADCDPNTQASCECPGY-I 432  
Db 793 RQRCDGFLQDPGEGNCVDINECTSLSEPCRGFCSCINTVGSYTCQRNPLI---CARGYHA 849  
QY 433 LDDGFICTDIDCENGFCSG---VCHNLPGTFECIC 466  
Db 850 SDDGAKCVDNBECGTGVRHRCGEGVQVCHNLPGSYRGDC 886  
RESULT 9  
A54105  
fibillin-2 precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 02-Aug-2002  
C:Accession: A54105; S17063; S31101  
R:Zhang, H.; Apfelroth, S.D.; Hu, W.; Davis, E.C.; Sanguinetti, C.; Bonadio, J.; Mecham, J. Cell Biol. 124, 855-863, 1994  
A:Title: Structure and expression of fibillin-2, a novel microfibrillar component preferentially expressed in bone  
A:Reference number: A54105; MUID:94165150; PMID:8120105  
A:Accession: A54105

A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-2918 <ZHA>  
A:Cross-references: GB:U03272  
R:Lee, B.; Godfrey, M.; Vitale, E.; Hori, H.; Mattei, M.G.; Sarfarazi, M.; Tsipouras, Nature 352, 330-334, 1991  
A:Title: Linkage of Marfan syndrome and a phenotypically related disorder to two different chromosomes  
A:Reference number: S17062; MUID:91304567; PMID:1852206  
A:Accession: S17063  
A:Molecule type: mRNA  
A:Residues: 752-1489, 1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P', 1928 <LEE>  
A:Cross-references: EMBL:X62009  
R:Milewicz, D.M.  
submitted to the EMBL Data Library, December 1992  
A:Reference number: S31101  
A:Accession: S31101  
A:Molecule type: mRNA  
A:Residues: 752-1407, 'R', 1409-1489, 1791, 'GS', 1794-1796, 'QLI', 1922-1923, 'LD', 1926, 'P', 1928 <LEE>  
A:Cross-references: EMBL:X62009  
C:Genetics:  
A:Gene: GDB:FBN2  
A:Cross-references: GDB:128122; OMIM:121050  
A:Map position: 5q23-5q31  
C:Superfamily: fibillin 1; EGF homology  
C:Keywords: extracellular protein  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:30-2918/Product: fibillin-2 #status predicted <MAT>  
F:1245-1280/Domain: EGF homology <EGF>  
F:1970-2013/Domain: EGF homology <EGF>  
Query Match 12.6%; Score 366; DB 2; Length 2918;  
Best Local Similarity 33.1%; Pred. No. 3.2e-14;  
Matches 101; Conservative 25; Mismatches 97; Indels 82; Gaps 17;  
QY 244 DCSVENGGCEHACNAIPGAPRCQCPAGAAQLADGRSCATASATQSCNDLCEHFCVNPND-- 301  
Db 1244 ECMINNGCDTQCTNSEGSEYSCSEGYALMPDRSCA-----DIDE--CENNPDC 1293  
QY 302 -----OPGYSYSCMCETGYRLAADQHRCEVDVDCILEPSPCP--QRCVNTQGGFECHCYP 353  
Db 1294 DGGQCTNTPGEYRCLCYDGFMA5MDMKTCIDVNECDLNSNICMFGECENTKGSFICHCQL 1353  
QY 354 NYDLVDGE--CYEVPDPC--FRANCEYOCQPLN-QTSYLCVCAEGFAPTPH-----EPHR 403  
Db 1354 GYSVAKGTGTGCTD-VDECEIGHANCDMHASCLNIPGSCSCREGW--IGNGIKCIDLDE 1410  
QY 404 COMFCNOTACPADCDPNTQAS--CECEGYILDDGFICTDIDE-----CENG----- 448  
Db 1411 CSNGTHQCSINAGC--VNTPGSYRCACSEGF--TGDGFTCSVDVDECAENINLCENGQCLNVP 1468  
QY 449 -----GF-----CSGVCHNLPGTFECICGPDLSALAHIG- 477  
Db 1469 GAYRCBCEMGFTPASDRSCQDIDEC5FQNICVSGTCNNLPGMHFHCICDDGYELDRGTGN 1528  
QY 478 -TDGD 481  
Db 1529 CTDDID 1533  
RESULT 10  
A47221  
fibillin 1 precursor - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 02-Jun-1995 #sequence\_revision 25-Apr-1997 #text\_change 02-Aug-2002  
C:Accession: A47221; I54355; S17064; I59574; S17062; S62111; A34198  
R:Corson, G.M.; Chalberg, S.C.; Dietz, H.C.; Charbonneau, N.L.; Sakai, L.Y. Genomics 17, 476-484, 1993  
A:Title: Fibillin binds calcium and is coded by cDNAs that reveal a multidomain structure  
A:Reference number: A47221; MUID:94010947; PMID:7691719  
A:Accession: A47221  
A:Molecule type: mRNA  
A:Residues: 1-337, 'T', 339-1029 <COR>  
A:Cross-references: GB:X63556



R:Duesterhoeft, A.; Lauber, J.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: 223035  
A:Accession: T46488  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-741 <AAA>  
A:Cross-references: EMBL:AL137638  
A:Experimental source: adult testis; clone DKF2p43J065  
C:Genetics:  
A:Note: DKF2p43J065.1

Query Match 12.2%; Score 357; DB 2; Length 741;  
Best Local Similarity 31.6%; Pred. No. 3.6e-14;  
Matches 86; Conservative 40; Mismatches 102; Indels 44; Gaps 15;  
Qy 245 CSVNGGCEHACNAIPGAPRCQCPAGAAQALQADGRSCTA-SATQSCNDLCEHFCVNPNDQP 303  
Db 68 CAMEHNCCEQLCVNVPSCVFCQCYSGVALAEDGRVAVDYCASENHGCEHCV-NAD-- 124  
Qy 304 GSYSCMCTGYRLAADQHRCEVDVDCILEPSPQPCQRCVNTGGFECHECHYPNYDL-VDGEC 362  
Db 125 GSYLQCQHEGFALNPDEKCTKIDYCASSNHGQCHECVNTDSDSYSCHLKGFTLNPKKT 184  
Qy 363 VEPVDPCE--RANCEYQCPQNTQSYLCVCAEGFAPTPH-----EPHRCQMF 408  
Db 185 CRRINYCALNPGEHECVNNEE-SYCYRCHRGYTLDPNGKTCGRVDHCAQDHCCEBOLC 243  
Qy 409 NOTACPADCDNTQAS--CECEGVILDDGF-ICTDDIDG--ENGFGSCVCHNLPTGF 462  
Db 244 -----LNTDSFVCQSEGLNEDLKTCSRVDYCLLSDHG--CEYSCVNMDRSF 291  
Qy 463 ECICGPDALARRHGTDCDSKVPD---GDG 491  
Db 292 ACQC-PEGHVLRSDGKTC--AKLDSALGDHG 320

RESULT 13  
A49457  
fibulin-2 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 03-May-1994 #sequence\_revision 03-May-1994 #text\_change 02-Aug-2002  
C:Accession: A49495; S74095  
R:Pan, T.C.; Sasaki, T.; Zhang, R.Z.; Faessler, R.; Timpl, R.; Chu, M.L.  
J. Cell Biol. 123, 1269-1277, 1993  
A:Title: Structure and expression of fibulin-2, a novel extracellular matrix protein with  
A:Reference number: A49457; MUID:94064787; PMID:8245130  
A:Accession: A49457  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1221 <PAN>  
A:Cross-references: GB:X75285; NID:g437046; PIDN:CAA53040.1; PID:g437047  
R:Sasaki, T.; Mann, K.; Murphy, G.; Chu, M.L.; Timpl, R.  
Eur. J. Biochem. 240, 427-434, 1996  
A:Title: Different susceptibilities of fibulin-1 and fibulin-2 to cleavage by matrix metalloproteinases  
A:Reference number: S74094; MUID:96439073; PMID:8841408  
A:Accession: S74095  
A:Molecule type: protein  
A:Residues: 236-238, 'X', 240-247; 260-275; 336-344, 'L', 346-361; 405-426; 566-568, 'EM', 569-589  
C:Superfamily: fibulin-2; EGF homology  
C:Keywords: calcium binding; duplication; extracellular matrix; glycoprotein; homotrimer  
F:1942-978/Domain: EGF homology <EG>

Query Match 12.2%; Score 357; DB 2; Length 1221;  
Best Local Similarity 29.5%; Pred. No. 5.4e-14;  
Matches 123; Conservative 39; Mismatches 141; Indels 114; Gaps 27;  
Qy 154 EQQCCEVAD-GFLCEFHPPATC-----RPLAV-----EPGAAAVS-----IT 192  
Db 521 EQQCESPNLGYPCN-HVMLSCCEBEPPLVPEVRRPPEAAPPVSEMAASREALS 579  
Qy 193 YGT-----PFAARGADFO--ALP-----VGSSAAVAPLGLQLM-----CTAPPG 229

Db 580 LGTEAELPNSLPGDDQDECLMLPGLCOHLICINTVSGSYRCACFPGLQDGRTRCPDRG 639  
Qy 230 AVQGHWARE-APGAWDCSV-----ENGCGEACNAIPGAPRCQCPAGAA 272  
Db 640 APOLDTAREAPRSASQVSPNTIPLVPQPNCKDNGPCQVGVGDTAMSCFPGYA 699  
Qy 273 LQADGRSC-----TASATQSCNDLCEHFCVNPNDQPGSYSC-----MCETGYRLAADQHR 322  
Db 700 IMADGVSCDEDDCLMGTDCS--WKQFCV---NTLGSFYCVNHTVLCAEGYILNA-HRK 753  
Qy 323 CEDVDDCILEPSPC--PORCVNTGGFECHECH-----CYNYDLVDGECVEPVDPCFRA--NC 374  
Db 754 CVDINECVTLHTCTRAEHCNVTPGSCQCYKALTCCEGYVLTGECTD-VDECVTGHNC 812  
Qy 375 E-YQCOPLNOTSYLCV-----CAEGFAPTPH-----EPHRCQMFNQTACPA 415  
Db 813 QAGFSCQN-TKGSFYCAQRCMDGFLQDPGNCVDINECTSLLEPCRSFSCINTVGSY 871  
Qy 416 DCDPNTQASCECEGY-ILDDGFTCTDIDCEENGFCGSG---VCHNLPGTFECICGP 468  
Db 872 TCQRNPLV---CGRGYHANEESGCVDNBCECTGVHRCGEGOLCYNLPGSYRCDCKP 925

RESULT 14  
T09059  
notch4 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 11-Jan-2000  
C:Accession: T09059  
R:Rowen, L.; Mahairs, G.; Qin, S.; Ahearn, M.E.; Dankers, C.; Loretz, C.;  
submitted to the EMBL Data Library, October 1997  
A:Description: Sequence of the mouse major histocompatibility locus class III region.  
A:Reference number: Z16543  
A:Accession: T09059  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1964 <ROW>  
A:Cross-references: EMBL:AF030001; NID:g2564945; PID:g2564947  
C:Genetics:  
A:Gene: notch4  
A:Map position: 17  
A:Introns: 22/1; 49/2; 148/1; 264/1; 305/1; 384/1; 436/1; 501/1; 539/1; 577/1; 618/1;  
1679/3; 1729/1; 1761/3  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
C:Keywords: receptor; signal transduction  
F:514-545/Domain: EGF homology <EGF>

Query Match 11.4%; Score 332.5; DB 2; Length 1964;  
Best Local Similarity 26.3%; Pred. No. 2.3e-12;  
Matches 128; Conservative 34; Mismatches 142; Indels 183; Gaps 31;  
Qy 134 GPLCVAVSAAEATVPSEPIWEEQCEVKADGFLCE-FHFPTATCRPLAVEPGAAAAAVSIT 192  
Db 35 GGTCLRLSRGGIC-----QC---APGLGETCQFPDPCR----- 66  
Qy 193 YGTFFAARGADFOAL---PVGSSAAVAPLGLQLMCTAPPG---AYOGHWAREAPGAWDC 245  
Db 67 -DTQLCKNGSGSCQALLTPPSRSRPTSLTPHFSCPTSGTSGTDRCTHLEELCPSPF-C 124  
Qy 246 SVENGCGEACNAIPGAPRCQCPAGAAQALQADGRSCTASATQSCNDLCEHFCVNPNDQP 304  
Db 125 S--NGG--HCYQASGRPQCCEPGWT-----GEOCQ-----LRDFCSANPCANGG 166

Qy 305 -----SYSCMCTGYRLAADQHRCE-DVDDCILEPSPCPQ---RCVNTGGFECHECH--- 351  
Db 167 VCLATYPIQICRCPPGF-----EGHTCERDINECFLEPGCPQGTSCHTLGSYQCLCPVG 222  
Qy 352 --YNYDLVDGECVEPVDPCFRANCEYQCPPL---NNTSYLCVCAEGFAPTPHE----- 400  
Db 223 QEGPQCKLRKAGAC--PPGSCILNGG---TCQLVPEGHSTFHLCLCPPGTGLDCEMNPDDC 277  
Qy 401 -PHRCQMFNQTACPADCDPNTQASCECEGYILDDGFICT-DIDECE-----NGGF 450

Wed Dec 18 11:06:40 2002

```

Db 278 VRHQQ-----NGATCLDGLDLYT-----CLCPKTIW-----KWDCSEIDICEAARGPPRCRNGGT 328
QY 451 C-----SGVCHNLPGTFECICGP-----468
Db 329 CONTRAGSHVCVSWGAGCEEMLDCAATCAPGTCIDRVGSFSLCPLCPGRTGLLCH 388
QY 469 --DSALAR--HIGTDCDSGKVDG-----GDSGS-----GEPSPSP-----499
Db 389 LEDMCLSQPCHVNAQCSNPLTGLICQPGYSGSTCHQDLDECQMAQQGSPCEHGG 448
QY 500 ---TPGS 503
Db 449 CINTPGS 455

RESULT 15
T43210
fibulin-1D precursor - Caenorhabditis elegans (fragment)
C:Species: Caenorhabditis elegans
C:Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 02-Sep-2000
C:Accession: T43210
R:Barth, J.L.; Argaves, K.M.; Roark, E.F.; Little, C.D.; Argaves, W.S.
submitted to the EMBL Data Library, June 1998
A:Description: Identification of chicken and C. elegans fibulin-1 homologs and character
A:Reference number: 222337
A:Accession: T43210
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-589 <BAR>
A:Cross-references: EMBL:AF070477; PIDN:AAC24035.1
C:Genetics:
A:Note: intron positions not resolved (incomplete sequence)
C:Superfamily: unassigned EGF-related proteins; EGF homology

Query Match 11.4% Score 331; DB 2; Length 589;
Best Local Similarity 28.1%; Pred. No. 1.1e-12; Indels 108; Gaps 28;
Matches 126; Conservative 49; Mismatches 166;

QY 103 GPLRGF-QWVTGDNNTSYSRWRLDNLGAPLCPLCVASAAEATVPSEPI-----WEEQ 156
Db 19 GCLRSFNKCCNGDIEITH--ASEIITGRPLNDPHVLHLGDRCASSHCHLCHDRGGEKV 75
QY 157 QCEVKA-----DGFLECEHFP--ATCRPLAVE-----PGAAAAAVSITYGTFFA--- 198
Db 76 ECSSRGFDLAPDGMACVDHIDECAITLMDCLSQRLNTPGSPFKCIRTLSCTGYAMDS 135
QY 199 --ARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQ-----GHWAREAPGAWDCS---- 246
Db 136 ETERCRDVECNLGSN----DCGPLYQCRNTQGSYRCDAKKCGDGLQNPMTGECTSITC 191
QY 247 -----VENGGGE-----HACNA-----IPGAPRCQ-----CPAGAALQADGRSCTA 282
Db 192 PNGYYPKMGMCNDIDECVTGHNGAGEECVNTPGSFRCQKGNLCAHGYEVN----- 243
QY 283 SATQSCNDL--CEH-----FCVNPDPQSGYSQWCEYGRLAADQHRCDVDDCIL--- 331
Db 244 GATGCEVDNECQQQVCGSMCEI---NLPGYKCKGPGYGFNDKAKCEDVDCEIKFAG 300
QY 332 EPSPCQRCVNTQGGFECGYPNYDLV--DGECEVFPDPCFR--ANCEYQQQPLNQTSYLC 388
Db 301 HVCDLASACINTIGSFECKKPGFQLASDGRRCEDVNECTTIGIAACEQKCVNI--PGSYQC 359
QY 389 VCAEGFAPIP-----HEPHRCQMF--NOTACPADCDPNTQAS--CECPGEGY-ILDDGFI 438
Db 360 ICDRGFALGPDGDKCEDIDECSIWAGSGNDLCMGGC-INTKGSYLQCPGPKYIQPDGRT 418
QY 439 CTDIDCEGNGGFCSG---VCHNLPGTFEC 464
Db 419 CYDVECAMGE-CAGSDKVCVNTLGSFKC 446

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:20:02 ; Search time 10.5 Seconds  
(without alignments)  
2038.265 Million cell updates/sec

Title: US-09-509-994-2  
Perfect score: 2916  
Sequence: 1 MLGVILVGLALAGLGFAP.....PSPTPGSTLTTPPAVGLVHSG 516

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2916	100.0	575	1 TRBM_HUMAN	P07204 homo sapien
2	1903	65.3	577	1 TRBM_MOUSE	P15306 mus musculus
3	1095.5	37.6	356	1 TRBM_BOVIN	P06579 bos taurus
4	587	20.1	652	1 CD93_HUMAN	O89103 homo sapien
5	547	18.8	644	1 CD93_MOUSE	O89103 mus musculus
6	534.5	18.3	643	1 CD93_RAT	Q96t61 rattus norv
7	374	12.8	2907	1 FBNI_MOUSE	Q61555 mus musculus
8	368	12.6	2871	1 FBNI_BOVIN	P98133 bos taurus
9	367.5	12.6	1184	1 FBNI_HUMAN	P98095 homo sapien
10	366	12.6	2871	1 FBNI_PIG	Q9tv36 sus scrofa
11	366	12.6	2911	1 FBNI_HUMAN	P35556 homo sapien
12	364	12.5	2871	1 FBNI_MOUSE	P35555 homo sapien
13	362	12.4	2871	1 FBNI_MOUSE	Q61554 mus musculus
14	357	12.2	1221	1 FBNI_MOUSE	P37889 mus musculus
15	356	12.2	956	1 MTN2_HUMAN	O00339 homo sapien
16	332.5	11.4	1964	1 NTC4_MOUSE	P31695 mus musculus
17	330.5	11.3	1712	1 LTB1_RAT	Q00918 rattus norv
18	326.5	11.2	956	1 MTN2_MOUSE	Q08746 mus musculus
19	324.5	11.1	712	1 LTB1_CAEEL	O77469 caenorhabdi
20	324.5	11.1	1394	1 LTB1_HUMAN	P22064 homo sapien
21	322.5	11.1	1595	1 LTB1_HUMAN	Q14766 homo sapien
22	321	11.0	2321	1 NTC3_HUMAN	Q9um47 homo sapien
23	315.5	10.8	443	1 FBNI_HUMAN	Q95967 homo sapien
24	311	10.7	703	1 FBNI_HUMAN	P23142 homo sapien
25	308.5	10.6	443	1 FBNI_HUMAN	O55058 cricetus
26	308.5	10.6	2003	1 NTC4_HUMAN	Q9ubx5 homo sapien
27	308	10.6	448	1 FBNI_HUMAN	Q9ubx5 homo sapien
28	308	10.6	705	1 FBNI_MOUSE	Q08879 mus musculus
29	307	10.5	684	1 FBNI_CHICK	O73775 gallus gall
30	306.5	10.5	443	1 FBNI_MOUSE	Q9wvj9 mus musculus
31	302	10.4	448	1 FBNI_MOUSE	Q9wvj9 mus musculus
32	297	10.2	2531	1 FBNI_MOUSE	O01705 mus musculus
33	296	10.2	448	1 FBNI_RAT	Q9wvh8 rattus norv

RESULT 1

TRBM_HUMAN	294	10.1	2319	1	NTC3_RAT	Q9rl72
ID	TRBM_HUMAN	STANDARD;	PRT;	575	AA.	mus musculus
AC	P07204;					Q35516
DT	01-APR-1988 (Rel. 07, Created)					Q12805
DT	01-FEB-1991 (Rel. 17, Last sequence update)					Q9w30
DT	15-JUN-2002 (Rel. 41, Last annotation update)					Q04721
DE	Thrombomodulin precursor (Fetomodulin) (TM) (CD141 antigen).					Q46530
GN	THBD OR THRM.					P07207
OS	Homo sapiens (Human).					O61982
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					Q07008
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					O35568
OX	NCBI_TaxID=9606;					P46531
RN	[1]					P01132
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=88004395; PubMed=2820710;					
RA	Suzuki K., Kusumoto H., Devashiki Y., Nishioka J., Maruyama I.,					
RA	Zushi M., Kawahara S., Honda G., Yamamoto S., Horiguchi S.,					
RT	"Structure and expression of human thrombomodulin, a thrombin					
RT	receptor on endothelium acting as a cofactor for protein C					
RT	activation."					
RL	EMBO J. 6:1891-1897(1987).					
RN	[2]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=88024950; PubMed=2822087;					
RA	Wen D., Dittman W.A., Ye R.D., Deaven L.L., Majerus P.W., Sadler J.E.;					
RT	"Human thrombomodulin: complete cDNA sequence and chromosome					
RT	localization of the gene."					
RL	Biochemistry 26:4350-4357(1987).					
RN	[3]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=87317665; PubMed=2819876;					
RA	Jackman R.W., Beeler D.L., Fritze L., Soff G., Rosenberg R.D.;					
RT	"Human thrombomodulin gene is intron depleted: nucleic acid sequences					
RT	of the cDNA and gene predict protein structure and suggest sites of					
RT	regulatory control."					
RL	Proc. Natl. Acad. Sci. U.S.A. 84:6425-6429(1987).					
RN	[4]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=88227901; PubMed=2836377;					
RA	Shirai T., Shiojiri S., Ito H., Yamamoto S., Kusumoto H.,					
RA	Devashiki Y., Maruyama I., Suzuki K.;					
RT	"Gene structure of human thrombomodulin, a cofactor for thrombin-					
RT	catalyzed activation of protein C."					
RL	J. Biochem. 103:281-285(1988).					
RN	[5]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=88227901; PubMed=2836377;					
RA	Shirai T., Shiojiri S., Ito H., Yamamoto S., Kusumoto H.,					
RA	Devashiki Y., Maruyama I., Suzuki K.;					
RT	"Gene structure of human thrombomodulin, a cofactor for thrombin-					
RT	catalyzed activation of protein C."					
RL	J. Biochem. 103:281-285(1988).					
RN	[5]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=21638749; PubMed=11780052;					
RA	Deloukas P., Matthews L.H., Ashurst J., Babbage A.K., Bagguley C.L.,					
RA	Jones M., Stavrides G., Almeida J.P., Beard L.M., Beare D.M.,					
RA	Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,					
RA	Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,					
RA	Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,					
RA	Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,					
RA	Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,					
RA	Coulson A., Coville G.J., Deadman R., Dhani P.D., Dunn M.,					

RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,  
RA Grahm D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,  
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,  
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,  
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,  
RA Lehtsalaho M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,  
RA Marsh V.L., Martin S.L., McConachie L.J., McEay K., McMurray A.A.,  
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,  
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,  
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,  
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkheen R., Sims S.,  
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,  
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,  
RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,  
RA Whitehead S.L., Whittaker P., Willey D.L., Williams L., Williams S.A.,  
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,  
RA Rogers J.;  
RT "The DNA sequence and comparative analysis of human chromosome 20";  
RL Nature 414:865-871(2001).  
RN [6]  
RP CARBOHYDRATE-LINKAGE SITE SER-492, AND MUTAGENESIS.  
RX MEDLINE-94029900; PubMed-8216207;  
RA Gerleitz B., Hassell T., Vlahos C.J., Parkinson J.F., Bang N.U.,  
RA Grinnell B.W.;  
RT "Identification of the predominant glycosaminoglycan-attachment site  
RT in soluble recombinant human thrombomodulin: potential regulation of  
RT functionality by glycosyltransferase competition for serine474";  
RL Biochem. J. 295:131-140(1993).  
RN [7]  
RP STRUCTURE BY NMR OF 389-407.  
RX MEDLINE-96007474; PubMed-7559494;  
RA Adler M., Seto M.H., Nitecki D.E., Lin J.H., Light D.R., Morser J.;  
RT "The structure of a 19-residue fragment from the C-loop of the fourth  
RT epidermal growth factor-like domain of thrombomodulin";  
RL J. Biol. Chem. 270:23366-23372(1995).  
RN [8]  
RP STRUCTURE BY NMR OF 364-407.  
RX MEDLINE-96100636; PubMed-8528067;  
RA Meiningen D.P., Hunter M.J., Komives E.A.;  
RT "Synthesis, activity, and preliminary structure of the fourth  
RT EGF-like domain of thrombomodulin";  
RL Protein Sci. 4:1683-1695(1995).  
RN [9]  
RP STRUCTURE BY NMR OF 427-444.  
RX MEDLINE-95034791; PubMed-7947766;  
RA Srinivasan J., Hu S., Hrabal R., Zhu Y., Komives E.A., Ni F.;  
RT "Thrombin-bound structure of an EGF subdomain from human  
RT thrombomodulin determined by transferred nuclear Overhauser  
RT effects";  
RL Biochemistry 33:13553-13560(1994).  
RN [10]  
RP STRUCTURE BY NMR OF 427-444.  
RX MEDLINE-96276211; PubMed-8745396;  
RA Hrabal R., Komives E.A., Ni F.;  
RT "Structural resiliency of an EGF-like subdomain bound to its target  
RT protein, thrombin";  
RL Protein Sci. 5:195-203(1996).  
RN [11]  
RP STRUCTURE BY NMR OF 405-444.  
RX MEDLINE-98035729; PubMed-9367781;  
RA Sampoli Benitez B.A., Hunter M.J., Meiningen D.P., Komives E.A.;  
RT "Structure of the fifth EGF-like domain of thrombomodulin: an  
RT EGF-like domain with a novel disulfide-bonding pattern";  
RL J. Mol. Biol. 273:913-926(1997).  
RN [12]  
RP VARIANT TED TYR-486.  
RX MEDLINE-95111115; PubMed-7811989;  
RA Oehlin A.-K., Marlar R.A.;  
RT "The first mutation identified in the thrombomodulin gene in a  
RT 45-year-old man presenting with thromboembolic disease";  
RL Blood 85:330-336(1995).  
RN [13]  
RP VARIANT TED Y-486, AND VARIANTS T-43; A-79; S-495 AND L-501.  
RX MEDLINE-97341986; PubMed-9198186;  
RA Oehlin A.-K., Norlund L., Marlar R.A.;  
RT "Thrombomodulin gene variations and thromboembolic disease";  
RL Thromb. Haemost. 78:396-400(1997).  
RN [14]  
RP VARIANT VAL-473.  
RX MEDLINE-97206518; PubMed-9157575;  
RA Norlund L., Holm J., Zoller B., Oehlin A.-K.;  
RT "A common thrombomodulin amino acid dimorphism is associated with  
RT myocardial infarction";  
RL Thromb. Haemost. 77:248-251(1997).  
RN [15]  
RP VARIANT THR-43.  
RX MEDLINE-99057299; PubMed-9843165;  
RA Doggen C.J.M., Kunz G., Rosendaal F.R., Lane D.A., Vos H.L.,  
RA Stubbs P.J., Manger C.V., Ireland H.;  
RT "A mutation in the thrombomodulin gene, 127G to A coding for Ala25Thr,  
RT and the risk of myocardial infarction in men";  
RL Thromb. Haemost. 80:743-748(1998).  
RN [16]  
RP VARIANT VAL-473.  
RX MEDLINE-21143723; PubMed-11245641;  
RA Wu K.K., Aleksic N., Ahn C., Boerwinkle E., Folsom A.R.,  
RA Juneja H.;  
RT "Thrombomodulin Ala455Val polymorphism and risk of coronary heart  
RT disease";  
RL Circulation 103:1386-1389(2001).  
CC -!- FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR  
CC THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS  
CC COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE  
CC ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA  
CC SCISSIONS THE ACTIVATED COPACTORS OF THE COAGULATION MECHANISM,  
CC FACTOR VA AND FACTOR VIIIa, AND THEREBY REDUCES THE AMOUNT OF  
CC THROMBIN GENERATED.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING  
CC THROMBOMODULIN.  
CC -!- POLYMORPHISM: VARIATIONS IN THBD ARE ASSOCIATED WITH AN INCREASED  
CC RISK OF DEVELOPING THROMBOEMBOLIC DISEASES (TED).  
CC -!- DISEASE: DEFECTS IN THBD COULD BE THE CAUSE OF INHERITED TED,  
CC ALSO CALLED INHERITED THROMBOPHILIA. PATIENTS WITH TED HAVE  
CC DEFECTS OF THE HAEMOPOIETIC SYSTEM WHICH CREATES A TENDENCY TO THE  
CC OCCURRENCE OF THROMBOSIS. TED PLAY AN IMPORTANT ROLE IN THE  
CC PATHOGENESIS OF VARIOUS CARDIOVASCULAR DISORDERS.  
CC -!- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.  
CC -!- DATABASE: NAME-PROW; NOTE-CD guide CD141 entry;  
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd141.htm".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X05495; CAA29045.1; -;  
CC EMBL; M16552; AAB59508.1; -;  
CC EMBL; J02973; AAA61175.1; -;  
CC EMBL; D00210; BAA00149.1; -;  
CC EMBL; AL049651; CAB51954.1; -;  
CC PIR; A27073; A27073.  
CC PIR; A28307; A28307.  
CC PIR; A29680; A29680.  
CC PDB; 1EGD; 15-NOV-95.  
CC PDB; 1FGE; 20-JUN-96.  
CC PDB; 1TMR; 08-JUN-95.  
CC PDB; 1ZAO; 29-JAN-96.  
CC PDB; 1ADX; 24-DEC-97.  
CC PDB; 2ADX; 24-DEC-97.  
CC GlycosuiteDB; P07204; -;  
CC Genew; HGNC:11784; THBD.



ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA  
 CC SCISSONS THE ACTIVATED CAPSOTACTORS OF THE COAGULATION MECHANISM,  
 CC FACTOR VA AND FACTOR VIIIA, AND THEREBY REDUCES THE AMOUNT OF  
 CC THROMBIN GENERATED.  
 CC  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC  
 CC -1- TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING  
 CC THROMBOMODULIN (BY SIMILARITY).  
 CC  
 CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.  
 CC  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC -----  
 CC EMBL: X14432; CAA32597.1; -  
 CC PIR: A32001; A32001.  
 CC PIR: S08488; S08488.  
 CC HSSP: P07204; IEGF.  
 CC  
 CC MGI: 98736; thbd.  
 CC  
 CC InterPro: IPR001152; ASX\_hydroxyl.  
 CC InterPro: IPR000561; EGF-like.  
 CC InterPro: IPR001881; EGF-CA.  
 CC InterPro: IPR001304; Lectin\_C.  
 CC InterPro: IPR001491; Thrombomodulin.  
 CC Pfam: PF00008; EGF; 4.  
 CC Pfam: PF00059; lectin\_c; 1.  
 CC PRINTS: PR00907; THROMBOMODULN.  
 CC SMART: SM00034; CLECF; 1.  
 CC SMART: SM00179; EGF\_CA; 1.  
 CC SMART: SM00001; EGF\_like; 5.  
 CC PROSITE: PS00010; ASX\_HYDROXYL; 2.  
 CC PROSITE: PS00022; EGF\_1; FALSE\_NEG.  
 CC PROSITE: PS01186; EGF\_2; 3.  
 CC PROSITE: PS01187; EGF\_CA; 2.  
 CC PROSITE: PS00441; C\_TYPE\_LECTIN\_2; 1.  
 CC Endothelial cell; Receptor; Blood coagulation; Repeat; Transmembrane;  
 CC Glycoprotein; Signal; EGF-like domain.  
 CC  
 CC SIGNAL 1 16  
 CC CHAIN 17 577  
 CC THROMBOMODULIN.  
 CC  
 CC DOMAIN 17 517  
 CC EXTRACELLULAR (POTENTIAL).  
 CC  
 CC TRANSMEM 518 541  
 CC POTENTIAL.  
 CC  
 CC DOMAIN 542 577  
 CC CYTOPLASMIC (POTENTIAL).  
 CC  
 CC DOMAIN 240 280  
 CC EGF-LIKE 1.  
 CC  
 CC DOMAIN 283 323  
 CC EGF-LIKE 2.  
 CC  
 CC DOMAIN 324 362  
 CC EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
 CC  
 CC DOMAIN 364 404  
 CC EGF-LIKE 4.  
 CC  
 CC DOMAIN 403 439  
 CC EGF-LIKE 5.  
 CC  
 CC DOMAIN 440 480  
 CC EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).  
 CC  
 CC DISULFID 244 255  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 251 264  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 266 279  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 287 295  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 291 307  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 309 322  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 328 339  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 345 348  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 350 361  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 368 377  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 373 387  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 389 403  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 407 416  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 412 424  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 426 438  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 444 454  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 449 463  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 465 479  
 CC BY SIMILARITY.  
 CC  
 CC CARBOHYD 113 113  
 CC N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC  
 CC CARBOHYD 243 243  
 CC N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC  
 CC CARBOHYD 256 256  
 CC N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC  
 CC CARBOHYD 408 408  
 CC N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC

FT	17	577	THROMBOMODULIN.
FT	17	517	EXTRACELLULAR (POTENTIAL).
FT	17	517	POTENTIAL.
FT	518	541	CYTOSOLIC (POTENTIAL).
FT	542	577	EGF-LIKE 1.
FT	240	280	EGF-LIKE 2.
FT	283	323	EGF-LIKE 3.
FT	324	362	EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT	364	404	EGF-LIKE 4.
FT	403	439	EGF-LIKE 5.
FT	440	480	EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT	244	255	BY SIMILARITY.
FT	251	264	BY SIMILARITY.
FT	266	279	BY SIMILARITY.
FT	287	295	BY SIMILARITY.
FT	291	307	BY SIMILARITY.
FT	309	322	BY SIMILARITY.
FT	328	339	BY SIMILARITY.
FT	335	348	BY SIMILARITY.
FT	350	361	BY SIMILARITY.
FT	368	377	BY SIMILARITY.
FT	373	387	BY SIMILARITY.
FT	389	403	BY SIMILARITY.
FT	407	416	BY SIMILARITY.
FT	412	424	BY SIMILARITY.
FT	426	438	BY SIMILARITY.
FT	444	454	BY SIMILARITY.
FT	449	463	BY SIMILARITY.
FT	465	479	BY SIMILARITY.
FT	113	113	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	243	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	256	256	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	408	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	408	408	N-LINKED (GLCNAC. . .) (POTENTIAL).



```
Db 123 CHCDTGVELVDECEVDVDFCDNCEYQCPVGRSEHKCAIEGAPVPGAPHKQCMFC 182
QY 409 NOTACPADCPNTQASCECEPEGYLDGFTCTDIDCEGNGFCGSGVCHNLPGTFEICGP 468
Db 183 NOTSCPADCPHYTCRCPEGYIIDEGSTCTDINECDT-NICPGQCHNLPGTYEICGP 241
QY 469 DSALARIHGTDCDSKV-----DGGSGSGEPPPTPGSTLTP-PA-VGLVHSG 516
Db 242 DSALSGQIGIDCDPTQVNEERGCTPDYG--GSGEPVPTPGATARPSPAPAGPLHSG 297

RESULT 4
CD93_HUMAN
ID Q9NPY3; O00274; STANDARD; PRT; 652 AA.
AC Q9NPY3; O00274;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Complement component C1q receptor precursor (Complement component 1, q
DE subcomponent, receptor 1) (ClqR) (ClqR(p)) (Clq/MBL/SPA receptor)
DE (CD93 antigen) (CDw93).
GN C1QR1 OR CD93.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX PubMed=9047234;
RA Nepomuceno R.R., Henschen-Edman A.H., Burgess W.H., Tenner A.J.;
RT "cDNA cloning and primary structure analysis of ClqR(p), the human
RT C1q/MBL/SPA receptor that mediates enhanced phagocytosis in vitro.";
RL Immunity 6:119-129(1997).
[2]
RP SEQUENCE FROM N.A., AND VARIANT ALA-318.
RX PubMed=11781389;
RA Steinberger P., Szekeres A., Wille S., Stockl J., Selenko N.,
RA Prager E., Staffler G., Madic O., Stockinger H., Knapp W.;
RT "Identification of human CD93 as the phagocytic C1q receptor (ClqR)
RT by expression cloning.";
RL J. Leukoc. Biol. 71:133-140(2002).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638749; PubMed=11780052;
RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baggeley C.L.,
RA Bailey J., Barlow K.F., Bates K.N., Bead L.M., Beare D.M.,
RA Beasley O.P., Bird C.P., Blakey S.E., Bridgman A.M., Brown A.J.,
RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
RA Clegg S., Copley V.E., Collier R.E., Connor R.E., Corby N.R.,
RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
RA Levaslahti M.H., Leversha M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
RA Marsh V.L., Martin S.L., McConachie L.J., McLay K., McMurray A.A.,
RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkeen R., Sims S.,
RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
RA Tracey A.C., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
RA Rogers J.;
RT "The DNA sequence and comparative analysis of human chromosome 20.";
RL Nature 414:865-871(2001).
[4]
```

```
RP SEQUENCE FROM N.A.
RC TISSUE=Leukocyte;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
[5]
RP CHARACTERIZATION.
RX PubMed=11994479;
RA McGreal E.P., Ikegaki N., Akatsu H., Morgan B.P., Gasque P.;
RT "Human C1qR is identical with CD93 and the mmi-11 antigen but does
RT not bind C1q.";
RL J. Immunol. 168:5222-5232(2002).
[6]
RP O-GLYCOSYLATION.
RX PubMed=10092817;
RA Nepomuceno R.R., Ruiz S., Park M., Tenner A.J.;
RT "ClqR is a heavily O-glycosylated cell surface protein involved in
RT the regulation of phagocytic activity.";
RL J. Immunol. 162:3583-3589(1999).
CC -1- FUNCTION: Receptor (or element of a larger receptor complex) for
CC C1q, mannose-binding lectin (MBL2) and pulmonary surfactant
CC protein A (SPA). May mediate the enhancement of phagocytosis in
CC monocytes and macrophages upon interaction with soluble defense
CC collagens. May play a role in intercellular adhesion.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: Highly expressed in endothelial cells,
CC platelets, cells of myeloid origin, such as monocytes and
CC neutrophils. Not expressed in cells of lymphoid origin.
CC -1- PTM: N- and O-glycosylated.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.
CC -1- CAUTION: Has been sometimes referred to as a collectin receptor.
CC -1- CAUTION: According to Ref.5, Clq is not a ligand for C1QR1.
CC -1- DATABASE: NAME=PROW; NOTE=PROW 3:1-6(2001); /467246456_g.htm".
CC WWW="http://www.ncbi.nlm.nih.gov/prow/guide/467246456_g.htm".
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; U94333; AAB53110.1; -
CC EMBL; AL118508; CAC00597.1; -
CC EMBL; BC028075; AAB28075.1; -
CC HSSP; P35555; 1EMN.
CC MIM; 120577; -
CC InterPro; IPR000152; ASX_hydroxyl.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR001881; EGF_CA.
CC InterPro; IPR001304; Lectin_C.
CC InterPro; IPR001187; Tissue_factor.
CC Pfam; PF00008; EGF; 5.
CC Pfam; PF00059; lectin_c; 1.
CC Pfam; PF01108; Tissue_fac; 1.
CC SMART; SM00034; CLECT; 1.
CC SMART; SM00179; EGF_CA; 3.
CC SMART; SM00001; EGF_like; 2.
CC PROSITE; PS00010; ASX_HYDROXYL; 3. FALSE_NEG.
CC PROSITE; PS00615; C-TYPE_LECTIN_1; 1.
CC PROSITE; PS00041; C-TYPE_LECTIN_2; 1.
CC PROSITE; PS01186; EGF_2; 3.
CC PROSITE; PS01187; EGF_CA; 3.
CC Repeat; EGF-like domain; Signal; Transmembrane; Glycoprotein;
CC Signal; Lectin; Polymorphism.
CC SIGNAL 1 21
CC CHAIN 22 652 COMPLEMENT COMPONENT C1Q RECEPTOR.
CC DOMAIN 24 580 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 581 601 POTENTIAL.
CC DOMAIN 602 652 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 32 174 C-TYPE LECTIN.
CC DOMAIN 260 301 EGF-LIKE 1.
CC FT
```

```
FT DOMAIN 302 344 EGF-LIKE 2.
FT DOMAIN 345 384 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 385 426 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 427 468 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 594 601 POLY-LEU.
FT DISULFID 264 275 BY SIMILARITY.
FT DISULFID 271 285 BY SIMILARITY.
FT DISULFID 287 300 BY SIMILARITY.
FT DISULFID 306 317 BY SIMILARITY.
FT DISULFID 311 328 BY SIMILARITY.
FT DISULFID 330 343 BY SIMILARITY.
FT DISULFID 349 358 BY SIMILARITY.
FT DISULFID 354 367 BY SIMILARITY.
FT DISULFID 369 383 BY SIMILARITY.
FT DISULFID 389 400 BY SIMILARITY.
FT DISULFID 396 409 BY SIMILARITY.
FT DISULFID 411 425 BY SIMILARITY.
FT DISULFID 431 443 BY SIMILARITY.
FT DISULFID 439 452 BY SIMILARITY.
FT DISULFID 454 467 BY SIMILARITY.
FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 318 318 V -> A.
FT CONFLICT 22 22 T -> V (IN AA SEQUENCE).
FT CONFLICT 36 36 C -> T (IN AA SEQUENCE).
FT CONFLICT 38 39 TA -> RI (IN AA SEQUENCE).
FT CONFLICT 155 155 G -> N (IN REF. 1).
FT CONFLICT 186 186 G -> A (IN AA SEQUENCE).
FT CONFLICT 492 492 S -> A (IN AA SEQUENCE).
FT CONFLICT 496 496 R -> Q (IN AA SEQUENCE).
FT CONFLICT 504 504 R -> G (IN AA SEQUENCE).
FT CONFLICT 541 541 P -> S (IN REF. 1).
SQ SEQUENCE 652 AA; 68560 MW; BECAOFECAS5FCAC2 CRC64;

Query Match
Best Local Similarity 20.1%; Score 587; DB 1; Length 652;
Matches 182; Conservative 55; Mismatches 209; Indels 124; Gaps 32;

QY 2 LGVLVLGALAA--GLGPPAPAEPPQGGSCQVHDCFPALYPGPATFLNASQICDGLRHL 59
DB 5 MGLLLLLLLLLLTPQAGTADTE----AVCVGTACVTAHSGKLSAAEAQHNCNGNGL 60

QY 60 MTRVSSVAAD---VISILLNGDGVGR--RLMTGLQLPG--CGDKPKRLGRLGFWMT 112
DB 61 AFVKSKEAQHVQRVLAQLLRREALTARMSKFWIQLQREKGLCDPLS--PLKGF5WVG 118

QY 113 GDNNTSYRWARLDLNGAPLCGLCA--VSAAEATVPSE--PIWEEQQC-----EVKAD 163
DB 119 GGEDTPYNWHKELRNSC--ISKRCVSLLDLSQPLPSRLPWSEGGPGSPGSGSNIE 176

QY 164 GPLCFEHPATCRPIAV--EPGAAAAVSVITVTPFAARGADFOALPVGSSAAVAPLGLQL 222
DB 177 GFVCKEFGKMRPLALGGPG-----QVTVTTPFTTSSLEAVPFAAANVA----- 224

QY 223 MC-TAPPAVQGHW---AREAPGAWD-----CSVENGGCEHAC--NAIRGAP 263
DB 225 -CGEGDKDETSHYFLCKEKAPDFVDMGSSGPLCVSPKRYCNFNNGCHQDFEGGDSF 283

QY 264 RQCPAGAAQADGRSCVATASQNDLCE--HFCVNPDPQPGSYSCMCTGYRLAADOH 321
DB 284 LGCRRPFRLLDLVTC--ASRNPSSSPCRGATCVLGP--HGKNTYCRCPQGYLDSSQL 341

QY 322 REEDVDDCILBSPCPORCVNTQGGFECHVYPNYDLVDGCEVPVDFCFRANCYQCOPL 381
DB 342 DCVDVDEC--QDSPCAQECVNTPGGFCRCWVG-----EPGGP-----GEGACQDV 386

QY 382 NQTSYLCVCAEGFAPIPHEPRHCQMFNCQNTACPADCPDNTQAS--CECPGYIL--DDGF 437
DB 387 DB-----CALGRSP-----CAQGC--TNTDGSFHCSECEGYVLAGEDGT 423

QY 438 ICTDIDEC--ENGFGSGVGNLHPTFCICGPDSSALARHITGTDGSKV-----D 486
DB 424 QCQDVDECVGGPLCDLSLCFNTQGSFHCGLPGLWLAAP--GVSCWTGMPVSLGPPSPGPD 482
```

```
QY 487 GDSGSGE-----PPSPPTGSGTLTPPA 509
DB 483 EEDKGEKGVTPRAATASPTRGPEGTAKA 512

RESULT 5
CD93_MOUSE
ID CD93_MOUSE STANDARD; PRT; 644 AA.
AC O89103;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Complement component clg receptor precursor (Complement component 1, q subcomponent, receptor 1) (ClqR) (ClqR(p)) (Clq/MBL/SPA receptor) (CD93 antigen) (Cell surface antigen AA4) (Lymphocyte antigen 68).
GN CLQRI OR CD93 OR CLQRP OR LY68 OR AA4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RX PubMed=11074255;
RA Kim T.S., Park M., Nepomuceno R.R., Palmirini G., Winokur S., Cotman C.A., Bengtsson U., Tenner A.J.;
RT "Characterization of the murine homolog of ClqR(p): identical cellular expression pattern, chromosomal location and functional activity of the human and murine ClqR(p).";
RL Mol. Immunol. 37:377-389(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukemia;
RX PubMed=10403644;
RA Petrenko O., Beavis A., Klaine M., Kittappa R., Godin I., Lemischka I.R.;
RT "The molecular characterization of the fetal stem cell marker AA4.";
RL Immunity 10:691-700(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV; TISSUE=Spleen, and Endothelial cells;
RX MEDLINE=99359842; PubMed=10430665;
RA Norsworthy P.J., Taylor P.R., Walport M.J., Botto M.;
RT "Cloning of the mouse homolog of the 126-kDa human Clq/MBL/SP-A receptor, ClqR(p).";
RL Mamm. Genome 10:789-793(1999).
CC -I- FUNCTION: Receptor (or element of a larger receptor complex) for Clq, mannose-binding lectin (MBL2) and pulmonary surfactant protein A (SPA). May mediate the enhancement of phagocytosis in monocytes and macrophages upon interaction with soluble defense collagens. May play a role in intercellular adhesion. Marker for early multipotent hematopoietic precursor cells. May play a role in cell-cell interactions during hematopoietic and vascular development.
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC -I- TISSUE SPECIFICITY: Expressed in lung, heart and bone marrow. Expressed at lower level in ovary, whole embryo and fetal liver. Not detected in brain, adult liver or thymus. Highly expressed in peritoneal cavity and bone marrow macrophages. Not detected in epithelial cells.
CC -I- DEVELOPMENTAL STAGE: First detectable in day 9 embryos, in the endocardium and vascular endothelium in the anterior part of the embryo. Expression in endothelial cells, initially restricted to aorta, omphalomesenteric and umbilical arteries, later extends to subcardinal veins, intersomitic arteries and perineural vessels. On day 10, detectable in the entire embryo.
CC -I- PTM: N- and O-glycosylated (By similarity).
CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -I- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL; AF074856; AAC63274.1; -;  
DR EMBL; AF081789; AAC62649.1; -;  
DR EMBL; AF099939; AAD47906.1; -;  
DR EMBL; AF099938; AAD47906.1; JOINED.  
DR HSSP; P35555; 1EMN.  
DR MGD; MGI:106664; Ly68.  
DR InterPro; IPR000152; ASX\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR001304; Lectin\_C.  
DR Pfam; PF00008; EGF; 5.  
DR Pfam; PF00059; lectin\_c; 1.  
DR SMART; SM00034; CLECT; 1.  
DR SMART; SM00179; EGF\_CA; 3.  
DR SMART; SM00001; EGF\_like; 2.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 3.  
DR PROSITE; PS00615; C\_TYPE\_LECTIN\_1; FALSE\_NEG.  
DR PROSITE; PS00041; C\_TYPE\_LECTIN\_2; 1.  
DR PROSITE; PS01186; EGF\_2; 3.  
DR PROSITE; PS01187; EGF\_CA; 3.  
KW Receptor; EGF-like domain; Signal; Transmembrane; Glycoprotein;  
KW Repeat; Lectin.  
FT SIGNAL 1 22 POTENTIAL.  
FT CHAIN 23 644 COMPLEMENT COMPONENT C1Q RECEPTOR.  
FT DOMAIN 23 572 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 573 593 POTENTIAL.  
FT DOMAIN 594 644 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 31 173 C-TYPE LECTIN.  
FT DOMAIN 257 298 EGF-LIKE 1.  
FT DOMAIN 299 341 EGF-LIKE 2.  
FT DOMAIN 342 381 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 382 423 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 424 465 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).  
FT DOMAIN 261 272 BY SIMILARITY.  
FT DISULFID 268 282 BY SIMILARITY.  
FT DISULFID 284 297 BY SIMILARITY.  
FT DISULFID 303 314 BY SIMILARITY.  
FT DISULFID 308 325 BY SIMILARITY.  
FT DISULFID 327 340 BY SIMILARITY.  
FT DISULFID 346 355 BY SIMILARITY.  
FT DISULFID 351 364 BY SIMILARITY.  
FT DISULFID 366 380 BY SIMILARITY.  
FT DISULFID 386 397 BY SIMILARITY.  
FT DISULFID 393 406 BY SIMILARITY.  
FT DISULFID 408 422 BY SIMILARITY.  
FT DISULFID 428 440 BY SIMILARITY.  
FT DISULFID 436 449 BY SIMILARITY.  
FT DISULFID 451 464 BY SIMILARITY.  
FT CARBOHYD 102 102 N-LINKED (GLCNAC... ) (POTENTIAL).  
FT CARBOHYD 322 322 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 644 AA; 69354 MW; EB4351648BF8635A CRC64;  
  
Query Match 18.8%; Score 547; DB 1; Length 644;  
Best Local Similarity 30.28; Pred. No. 4.3e-29;  
Matches 168; Conservative 55; Mismatches 210; Indels 124; Gaps 27;  
  
OY 5 LVLAGALAGLGFAPAPBQPGGSCQVHDCFALYPGPATFLNASQICDGLRGLHMTVRS 64  
DB 9 LLLGLLGPWAG--AAADSQ--AVVCEGTACTAHWGLSAAEAQHRGNGNGNLATVRS 64  
OY 65 SVAA-----DVISLLNGDGGVGR--RLWIGLQLPPGGCDPRGLGRGFQVWTDGNTWS 118  
DB 65 EEEARHVQALQTLKTRAPLEAKMKFWIGLQREKGNCTYHDL-PMRGFSVWVGGEATFA 123  
OY 119 YSRWRLDNLGAPICGCLPCVAVSAEAVYVSE-PTWEEQQCEV-----KADGFLCEHFH 171  
DB 124 YSNWYKASKSCIFRCVSLILDLSLTPHPSHLPKWHSPCGTPEAPGNSIEGFLCKFNF 183

OY 172 PATCRPLAV-EPGAAAAAVSITYGTFFAARGADFOALPVGSSAAVAPLGLQ-----LM 223  
DB 184 KGMRPLALGPG-----RVITYTPQATTSLEAVPFASVANVA-CGDEAKSETHYFL 236  
OY 224 CT-APPGAVOGHWAREAP-----GAWDCSVENGCGEHAC-NAIPGAPRCQCAGAAALQADG 277  
DB 237 CNEKTPGIF--HMGSSGPLCVSPKFGCSFNNGGQQDQCFEGGDSFGRCGCRPGRLLDDL 294  
OY 278 RSTASATQSCNDLCEHFCVNPNDPG-----SYSCMETGYRLAADOHRCEVDV 327  
DB 295 VTCAS-----RNPCSSNPCTGGGMCHSVLSENYTCRCPSGYQLDSSQVHCVDID 344  
OY 328 DCILSPSPORCVNTGGFECCHYPNYDLVDG---ECVEVPDPCFRANCEVCQCPQAT 384  
DB 345 EC--QDSPCAQDCVNTLGSFHCCEWVGQ--PSGPKKEACEDVDECAANSP----- 392  
OY 385 SYLCVCAEGFAPIPHEPHRCQMFNCNOTACPADCPDPTQAS--CECEGYIL--DDGFICT 440  
DB 393 -----CAQGGCI-----NTDGSFVCSCKEYIYSGEDSTQCE 423  
OY 441 DIDECEN--GGFCSGVCHNLPGTFECICGPDPSALARHGTDCDSG-----KVD 486  
DB 424 DIDECDARGNPDCLSCFTDGSFRGCGPPGWELAPN-GVFCRSRTVSELPARPPQKED 482  
OY 487 GDSGSGEPSPSPGSGS 503  
DB 483 NDRKESTMPTTEMPSS 499  
  
RESULT 6  
CD93\_RAT  
ID CD93\_RAT STANDARD; PRT; 643 AA.  
AC O9FT61; O9JI26;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Complement component C1q receptor precursor (Complement component 1, q subcomponent, receptor 1) (C1qR) (C1qR(p)) (C1q/MBL/SPA receptor)  
DE (CD93 antigen) (Cell surface antigen AA4).  
DE C1QR1 OR CD93 OR C1QRP.  
GN Rattus norvegicus (Rat).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-PVG; TISSUE-Natural killer cells;  
RX MEDLINE=20545218; PubMed=11093152;  
RA Lovik G., Vaage J.T., Dissen E., Szpirer C., Ryan J.C., Rolstad B.;  
RT "Characterization and molecular cloning of rat C1qR, a receptor on NK cells.";  
RL Eur. J. Immunol. 30:3355-3362(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Wistar; TISSUE-Lung;  
RX MEDLINE=20507883; PubMed=10934210;  
RA Dean Y.D., McGreal E.P., Akatsu H., Gasque P.;  
RT "Molecular and cellular properties of the rat AA4 antigen, a C-type lectin-like receptor with structural homology to thrombomodulin.";  
RL J. Biol. Chem. 275:34382-34392(2000).  
CC -!- FUNCTION: Receptor (or element of a larger receptor complex) for C1q, mannose-binding lectin (MBL2) and pulmonary surfactant protein A (SPA). May mediate the enhancement of phagocytosis in monocytes and macrophages upon interaction with soluble defense collagens. May play a role in intercellular adhesion.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: Widely expressed. Highly expressed in lung and heart. Expressed at lower level in brain, thymus, liver, spleen, intestine, kidney, adrenal gland, muscle and testis. Expressed on endothelial cells, platelets, undifferentiated monocytes and circulating natural killer cells.  
CC -!- PTM: N- and O-glycosylated (By similarity).



EGF-LIKE 42, CALCIUM-BINDING.  
EGF-LIKE 43, CALCIUM-BINDING.  
EGF-LIKE 44, CALCIUM-BINDING.  
EGF-LIKE 45, CALCIUM-BINDING.  
EGF-LIKE 46, CALCIUM-BINDING.  
EGF-LIKE 47, CALCIUM-BINDING.

2524	EGF-LIKE 42, C	BY SIMILARITY.
2563	EGF-LIKE 43, C	BY SIMILARITY.
2606	EGF-LIKE 44, C	BY SIMILARITY.
2646	EGF-LIKE 45, C	BY SIMILARITY.
2687	EGF-LIKE 46, C	BY SIMILARITY.
2727	EGF-LIKE 47, C	BY SIMILARITY.
120	BY SIMILARITY.	BY SIMILARITY.
130	BY SIMILARITY.	BY SIMILARITY.
141	BY SIMILARITY.	BY SIMILARITY.
159	BY SIMILARITY.	BY SIMILARITY.
164	BY SIMILARITY.	BY SIMILARITY.
175	BY SIMILARITY.	BY SIMILARITY.
190	BY SIMILARITY.	BY SIMILARITY.
196	BY SIMILARITY.	BY SIMILARITY.
207	BY SIMILARITY.	BY SIMILARITY.
292	BY SIMILARITY.	BY SIMILARITY.
301	BY SIMILARITY.	BY SIMILARITY.
316	BY SIMILARITY.	BY SIMILARITY.
334	BY SIMILARITY.	BY SIMILARITY.
343	BY SIMILARITY.	BY SIMILARITY.
358	BY SIMILARITY.	BY SIMILARITY.
503	BY SIMILARITY.	BY SIMILARITY.
512	BY SIMILARITY.	BY SIMILARITY.
526	BY SIMILARITY.	BY SIMILARITY.
542	BY SIMILARITY.	BY SIMILARITY.
551	BY SIMILARITY.	BY SIMILARITY.
566	BY SIMILARITY.	BY SIMILARITY.
584	BY SIMILARITY.	BY SIMILARITY.
593	BY SIMILARITY.	BY SIMILARITY.
608	BY SIMILARITY.	BY SIMILARITY.
625	BY SIMILARITY.	BY SIMILARITY.
634	BY SIMILARITY.	BY SIMILARITY.
649	BY SIMILARITY.	BY SIMILARITY.
666	BY SIMILARITY.	BY SIMILARITY.
675	BY SIMILARITY.	BY SIMILARITY.
690	BY SIMILARITY.	BY SIMILARITY.
777	BY SIMILARITY.	BY SIMILARITY.
786	BY SIMILARITY.	BY SIMILARITY.
801	BY SIMILARITY.	BY SIMILARITY.
819	BY SIMILARITY.	BY SIMILARITY.
828	BY SIMILARITY.	BY SIMILARITY.
843	BY SIMILARITY.	BY SIMILARITY.
859	BY SIMILARITY.	BY SIMILARITY.
868	BY SIMILARITY.	BY SIMILARITY.
883	BY SIMILARITY.	BY SIMILARITY.
964	BY SIMILARITY.	BY SIMILARITY.
973	BY SIMILARITY.	BY SIMILARITY.
988	BY SIMILARITY.	BY SIMILARITY.
1082	BY SIMILARITY.	BY SIMILARITY.
1091	BY SIMILARITY.	BY SIMILARITY.
1106	BY SIMILARITY.	BY SIMILARITY.
1124	BY SIMILARITY.	BY SIMILARITY.
1133	BY SIMILARITY.	BY SIMILARITY.
1149	BY SIMILARITY.	BY SIMILARITY.
1167	BY SIMILARITY.	BY SIMILARITY.
1176	BY SIMILARITY.	BY SIMILARITY.
1191	BY SIMILARITY.	BY SIMILARITY.
1209	BY SIMILARITY.	BY SIMILARITY.
1218	BY SIMILARITY.	BY SIMILARITY.
1233	BY SIMILARITY.	BY SIMILARITY.
1250	BY SIMILARITY.	BY SIMILARITY.
1259	BY SIMILARITY.	BY SIMILARITY.
1274	BY SIMILARITY.	BY SIMILARITY.
1292	BY SIMILARITY.	BY SIMILARITY.
1301	BY SIMILARITY.	BY SIMILARITY.
1316	BY SIMILARITY.	BY SIMILARITY.
1334	BY SIMILARITY.	BY SIMILARITY.
1343	BY SIMILARITY.	BY SIMILARITY.
1358	BY SIMILARITY.	BY SIMILARITY.
1377	BY SIMILARITY.	BY SIMILARITY.
1386	BY SIMILARITY.	BY SIMILARITY.
1399	BY SIMILARITY.	BY SIMILARITY.
1418	BY SIMILARITY.	BY SIMILARITY.

```

FT DISULFID 1412 1427 BY SIMILARITY.
FT DISULFID 1429 1440 BY SIMILARITY.
FT DISULFID 1446 1458 BY SIMILARITY.
FT DISULFID 1453 1467 BY SIMILARITY.
FT DISULFID 1469 1482 BY SIMILARITY.
FT DISULFID 1488 1499 BY SIMILARITY.
FT DISULFID 1494 1508 BY SIMILARITY.
FT DISULFID 1510 1523 BY SIMILARITY.
FT DISULFID 1529 1540 BY SIMILARITY.
FT DISULFID 1535 1549 BY SIMILARITY.
FT DISULFID 1551 1564 BY SIMILARITY.
FT DISULFID 1647 1659 BY SIMILARITY.
FT DISULFID 1654 1668 BY SIMILARITY.
FT DISULFID 1670 1683 BY SIMILARITY.
FT DISULFID 1689 1701 BY SIMILARITY.
FT DISULFID 1696 1710 BY SIMILARITY.
FT DISULFID 1712 1725 BY SIMILARITY.

Query Match 12.8%; Score 374; DB 1; Length 2907;
Best Local Similarity 32.9%; Pred. No. 4e-17;
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;

Qy 244 DCSVGGCGHACNAIPGACRQCPAGAAQADGRSCTASATQSCNDLCEHFQVNPDP-- 301
Db 1238 ECTMNGGCTDQCTNSGSEYECSEGYALMPDGRSCA-----DIDE--CENPDIC 1287

Qy 302 -----QPGSYSCMCTGYRLAADQRHCRVDVDCILEPSPCP-QRCVNTQGGFECHCYP 353
Db 1288 DGGCTNIPGEYRCLVDFGFMASDMKTCIDVNECDLNPNCMFGECENTKGSFICHCQL 1347

Qy 354 NYDLVDGE--CVEPVDPQ--FRANCEVQCOPLN-QSYLCVCAEFGA-----PVP 398
Db 1348 GYSVKKTGTGCTD-VDECEIHAHNDHMSCLNVPVGSFKCSREGVWNGIKCIDIDDECA 1406

Qy 399 HEPHRCQMFQACADCPDQNTQAS--CCEPGYILDDGFICTDIDE-----CENG- 448
Db 1407 NGTHQCSI-----NAQC-VNTPGYSYRCACSEGF-TGDGFTCSVDDECAENTNLCENGQ 1457

Qy 449 -----GF-----CS-----GVCNLPFTFECICGPDPSALA 473
Db 1458 CLNVPGAYRCECEMGFTPASDSRSCQIDECSPQICVFCGTCNNLPMPFHCICDDGYELD 1517

Qy 474 RHIG--TDGD 481
Db 1518 RTGNCNTDID 1527

RESULT 8
ID FBN1_BOVIN STANDARD; PRT; 2871 AA.
AC P98133;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibrillin 1 precursor (MP340).
GN FBN1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=95137597; PubMed=7835900;
RA Tilstra D.J., Potter K.A., Byers P.H.;
RT "Sequencing of the coding region of the bovine fibrillin cDNA and
RL localization to bovine chromosome 10.";
RN Genomics 23:480-485(1994).
RN [2]
RP PARTIAL SEQUENCE.
RX MEDLINE=96132851; PubMed=8557636;
RA Gibson M.A., Hatzinikolas G., Kumaratilake J.S., Sandberg L.B.,

```

```

RA Nicholl J.K., Sutherland G.R., Cleary E.G.;
RT "Further characterization of proteins associated with elastic fiber
RL microfibrils including the molecular cloning of MAGP-2 (MP25).";
RL J. Biol. Chem. 271:1096-1103(1996).
CC -|- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS
CC THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE
CC LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.
CC -|- PFM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER
CC FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE
CC MICROFIBRILS.
CC -|- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.
CC -|- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L28748; AAA74122.1; -.
DR HSP; P35555; IAPJ.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF-II.
DR InterPro; IPR002212; Fibrill-assoc.
DR Pfam; PF00008; EGF; 46.
DR Pfam; PF00683; TB; 9.
DR PRINTS; PR00010; EGFBL00D.
DR SMART; SM00179; EGF_CA; 42.
DR SMART; SM00001; EGF_Like; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 43.
DR PROSITE; PS00022; EGF_L; 2.
DR PROSITE; PS01186; EGF_2; 38.
DR PROSITE; PS01187; EGF_CA; 43.
KW Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW Repeat; Signal; Multigene family.
FT SIGNAL 1 27
FT CHAIN 28 2871
FT FIBRILLIN 1.
FT EGF-LIKE 1.
FT EGF-LIKE 2.
FT EGF-LIKE 3.
FT EGF-LIKE 4.
FT EGF-LIKE 5.
FT TGFBP 1.
FT TGFBP 2.
FT TGFBP 3.
FT EGF-LIKE 6.
FT EGF-LIKE 7.
FT EGF-LIKE 8.
FT EGF-LIKE 9.
FT EGF-LIKE 10.
FT EGF-LIKE 11.
FT EGF-LIKE 12.
FT EGF-LIKE 13.
FT EGF-LIKE 14.
FT EGF-LIKE 15.
FT EGF-LIKE 16.
FT EGF-LIKE 17.
FT EGF-LIKE 18.
FT EGF-LIKE 19.
FT EGF-LIKE 20.
FT EGF-LIKE 21.
FT EGF-LIKE 22.
FT EGF-LIKE 23.
FT EGF-LIKE 24.
FT EGF-LIKE 25.
FT EGF-LIKE 26.
FT TGFBP 4.
FT EGF-LIKE 27.

```



FT	DOMAIN	1648	1688	EGF-LIKE 28, CALCIUM-BINDING.	FT	DISULFID	1140	1153	BY SIMILARITY.
FT	REPEAT	1689	1765	TGFBB 5.	FT	DISULFID	1159	1171	BY SIMILARITY.
FT	DOMAIN	1766	1807	EGF-LIKE 29, CALCIUM-BINDING.	FT	DISULFID	1180	1180	BY SIMILARITY.
FT	DOMAIN	1808	1848	EGF-LIKE 30, CALCIUM-BINDING.	FT	DISULFID	1182	1195	BY SIMILARITY.
FT	DOMAIN	1849	1890	EGF-LIKE 31, CALCIUM-BINDING.	FT	DISULFID	1201	1212	BY SIMILARITY.
FT	DOMAIN	1891	1929	EGF-LIKE 32, CALCIUM-BINDING.	FT	DISULFID	1208	1221	BY SIMILARITY.
FT	DOMAIN	1930	1972	EGF-LIKE 33, CALCIUM-BINDING.	FT	DISULFID	1233	1236	BY SIMILARITY.
FT	DOMAIN	1973	2012	EGF-LIKE 34, CALCIUM-BINDING.	FT	DISULFID	1242	1254	BY SIMILARITY.
FT	DOMAIN	2013	2054	EGF-LIKE 35, CALCIUM-BINDING.	FT	DISULFID	1249	1263	BY SIMILARITY.
FT	REPEAT	2055	2126	TGFBB 6.	FT	DISULFID	1265	1278	BY SIMILARITY.
FT	DOMAIN	2127	2165	EGF-LIKE 36, CALCIUM-BINDING.	FT	DISULFID	1284	1296	BY SIMILARITY.
FT	DOMAIN	2166	2205	EGF-LIKE 37, CALCIUM-BINDING.	FT	DISULFID	1291	1305	BY SIMILARITY.
FT	DOMAIN	2206	2246	EGF-LIKE 38, CALCIUM-BINDING.	FT	DISULFID	1307	1320	BY SIMILARITY.
FT	DOMAIN	2247	2290	EGF-LIKE 39, CALCIUM-BINDING.	FT	DISULFID	1326	1339	BY SIMILARITY.
FT	DOMAIN	2291	2332	EGF-LIKE 40, CALCIUM-BINDING.	FT	DISULFID	1333	1348	BY SIMILARITY.
FT	REPEAT	2333	2401	TGFBB 7.	FT	DISULFID	1350	1361	BY SIMILARITY.
FT	DOMAIN	2402	2443	EGF-LIKE 41, CALCIUM-BINDING.	FT	DISULFID	1367	1380	BY SIMILARITY.
FT	DOMAIN	2444	2484	EGF-LIKE 42, CALCIUM-BINDING.	FT	DISULFID	1374	1389	BY SIMILARITY.
FT	DOMAIN	2485	2523	EGF-LIKE 43, CALCIUM-BINDING.	FT	DISULFID	1391	1402	BY SIMILARITY.
FT	DOMAIN	2524	2566	EGF-LIKE 44, CALCIUM-BINDING.	FT	DISULFID	1408	1420	BY SIMILARITY.
FT	DOMAIN	2567	2606	EGF-LIKE 45, CALCIUM-BINDING.	FT	DISULFID	1415	1429	BY SIMILARITY.
FT	DOMAIN	2607	2647	EGF-LIKE 46, CALCIUM-BINDING.	FT	DISULFID	1431	1444	BY SIMILARITY.
FT	DOMAIN	2648	2687	EGF-LIKE 47, CALCIUM-BINDING.	FT	DISULFID	1450	1461	BY SIMILARITY.
FT	DISULFID	85	94	BY SIMILARITY.	FT	DISULFID	1456	1470	BY SIMILARITY.
FT	DISULFID	89	100	BY SIMILARITY.	FT	DISULFID	1472	1485	BY SIMILARITY.
FT	DISULFID	102	111	BY SIMILARITY.	FT	DISULFID	1491	1502	BY SIMILARITY.
FT	DISULFID	119	129	BY SIMILARITY.	FT	DISULFID	1497	1511	BY SIMILARITY.
FT	DISULFID	123	134	BY SIMILARITY.	FT	DISULFID	1513	1526	BY SIMILARITY.
FT	DISULFID	136	145	BY SIMILARITY.	FT	DISULFID	1610	1622	BY SIMILARITY.
FT	DISULFID	150	160	BY SIMILARITY.	FT	DISULFID	1617	1631	BY SIMILARITY.
FT	DISULFID	154	166	BY SIMILARITY.					
FT	DISULFID	168	177	BY SIMILARITY.					
FT	DISULFID	250	262	BY SIMILARITY.					
FT	DISULFID	257	271	BY SIMILARITY.					
FT	DISULFID	273	286	BY SIMILARITY.					
FT	DISULFID	292	304	BY SIMILARITY.					
FT	DISULFID	299	313	BY SIMILARITY.					
FT	DISULFID	315	328	BY SIMILARITY.					
FT	DISULFID	453	465	BY SIMILARITY.					
FT	DISULFID	460	474	BY SIMILARITY.					
FT	DISULFID	476	488	BY SIMILARITY.					
FT	DISULFID	494	504	BY SIMILARITY.					
FT	DISULFID	499	513	BY SIMILARITY.					
FT	DISULFID	515	528	BY SIMILARITY.					
FT	DISULFID	534	546	BY SIMILARITY.					
FT	DISULFID	541	555	BY SIMILARITY.					
FT	DISULFID	557	570	BY SIMILARITY.					
FT	DISULFID	576	587	BY SIMILARITY.					
FT	DISULFID	582	596	BY SIMILARITY.					
FT	DISULFID	598	611	BY SIMILARITY.					
FT	DISULFID	617	628	BY SIMILARITY.					
FT	DISULFID	623	637	BY SIMILARITY.					
FT	DISULFID	639	652	BY SIMILARITY.					
FT	DISULFID	727	739	BY SIMILARITY.					
FT	DISULFID	734	748	BY SIMILARITY.					
FT	DISULFID	750	763	BY SIMILARITY.					
FT	DISULFID	769	781	BY SIMILARITY.					
FT	DISULFID	776	790	BY SIMILARITY.					
FT	DISULFID	792	805	BY SIMILARITY.					
FT	DISULFID	811	821	BY SIMILARITY.					
FT	DISULFID	816	830	BY SIMILARITY.					
FT	DISULFID	832	845	BY SIMILARITY.					
FT	DISULFID	914	926	BY SIMILARITY.					
FT	DISULFID	921	935	BY SIMILARITY.					
FT	DISULFID	937	950	BY SIMILARITY.					
FT	DISULFID	1032	1044	BY SIMILARITY.					
FT	DISULFID	1039	1053	BY SIMILARITY.					
FT	DISULFID	1055	1068	BY SIMILARITY.					
FT	DISULFID	1074	1086	BY SIMILARITY.					
FT	DISULFID	1081	1095	BY SIMILARITY.					
FT	DISULFID	1097	1111	BY SIMILARITY.					
FT	DISULFID	1117	1129	BY SIMILARITY.					
FT	DISULFID	1124	1138	BY SIMILARITY.					

Query Match 12.6%; Score 368; DB 1; Length 2871;  
Best Local Similarity 31.5%; Pred. No. 9.9e-17;  
Matches 111; Conservative 26; Mismatches 109; Indels 106; Gaps 20;

Qy 244 DCSVENGGCEHACNAIPGAPRCQCPAGAAQADQRSCSTASATQSCNDLCEHFCVNP--- 300  
Db 1200 ECSIMNGGCEFTFNSEGSYECSCQPGFALPQDRSCT-----DIDE--CEDNPNC 1249

Qy 301 -----DQPGSYSCMETGYRLAADQHRCEVDVDCILEPSPC-PQRCVNTGGFECHCYP 353  
Db 1250 DGGQCTNIPGYRCLYDGFMASEDMKTCVDVNECDLNPICLSGTCENTKSGFICHCDM 1309

Qy 354 NYDLVDGE--CVEPVDPCC--FRANCEYOCQPLNOT-SYLCVCACGFA-----PIP 398  
Db 1310 GYSGKKGTGCTD-INECEIGHNCDRHAVCNTAGSKCSFGWIGDKICTDLDECS 1368

Qy 399 HEPHRCQMFNQTCAPADCDPNTQAS--CECPEGYILDGFCITDIDEC-EN----- 447  
Db 1369 NGTHMCSQH-----ADC-KNTMGSYRCLCKEY-TGDGFTCTDLDECSNLCNGNQ 1419

Qy 448 -----GGF-----CS-----GVCHNLPGTFECICGPDSSALA 473  
Db 1420 CLNAPGGYRCBDMGFVPSADGKACEDIDCSLNPICVFGTCHNLPLGFRCECEIGVELD 1479

Qy 474 RHIG-----TDCDSK-VDDGDSGSGEPPTPGSTLTTPPAVLV 513  
Db 1480 RSGGNCVDNECLDPTTTCISGNCVNTPGSYTCDCCPD---FELNPTRVGCV 1527

RESULT 9  
FBL2\_HUMAN  
ID FBL2\_HUMAN STANDARD; PRT; 1184 AA.  
AC P98095;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Fibulin-2 precursor.  
GN FBLN2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;



DT	16-OCT-2001 (Rel. 40, Created)	FT	DOMAIN	1028	1069	EGF-LIKE 15, CALCIUM-BINDING.
DT	16-OCT-2001 (Rel. 40, Last sequence update)	FT	DOMAIN	1070	1112	EGF-LIKE 16, CALCIUM-BINDING.
DT	15-JUN-2002 (Rel. 41, last annotation update)	FT	DOMAIN	1113	1154	EGF-LIKE 17, CALCIUM-BINDING.
DE	Fibrillin 1 precursor.	FT	DOMAIN	1155	1196	EGF-LIKE 18, CALCIUM-BINDING.
GN	FN1.	FT	DOMAIN	1197	1237	EGF-LIKE 19, CALCIUM-BINDING.
OS	Sus scrofa (Pig).	FT	DOMAIN	1238	1279	EGF-LIKE 20, CALCIUM-BINDING.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	FT	DOMAIN	1280	1321	EGF-LIKE 21, CALCIUM-BINDING.
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.	FT	DOMAIN	1322	1362	EGF-LIKE 22, CALCIUM-BINDING.
OX	NCBI_TaxID=9823;	FT	DOMAIN	1363	1403	EGF-LIKE 23, CALCIUM-BINDING.
FN	[1]	FT	DOMAIN	1404	1445	EGF-LIKE 24, CALCIUM-BINDING.
RN	SEQUENCE FROM N.A.	FT	DOMAIN	1446	1486	EGF-LIKE 25, CALCIUM-BINDING.
RC	TISSUE=Lung;	FT	DOMAIN	1487	1527	EGF-LIKE 26, CALCIUM-BINDING.
RC	MEDLINE=99156858; PubMed=10036187;	FT	REPEAT	1528	1605	TGFBBP 4.
RA	Biery N.J.; Eldadah Z.A.; Moore C.S.; Stetten G.; Spencer F.,	FT	DOMAIN	1606	1647	EGF-LIKE 27, CALCIUM-BINDING.
RA	Dietz H.C.;	FT	DOMAIN	1648	1688	EGF-LIKE 28, CALCIUM-BINDING.
FT	gene expression."	FT	REPEAT	1689	1765	TGFBBP 5.
RL	Genomics 56:70-77(1999).	FT	DOMAIN	1766	1807	EGF-LIKE 29, CALCIUM-BINDING.
CC	-/- FUNCTION: STRUCTURAL COMPONENT OF CONNECTIVE TISSUE MICROFIBRILS	FT	DOMAIN	1808	1848	EGF-LIKE 30, CALCIUM-BINDING.
CC	THAT BINDS CALCIUM. FIBRILLIN-1-CONTAINING MICROFIBRILS PROVIDE	FT	DOMAIN	1849	1890	EGF-LIKE 31, CALCIUM-BINDING.
CC	LONG-TERM FORCE BEARING STRUCTURAL SUPPORT.	FT	DOMAIN	1891	1929	EGF-LIKE 32, CALCIUM-BINDING.
CC	-/- FTM: FORMS INTERMOLECULAR DISULFIDE BONDS EITHER WITH OTHER	FT	DOMAIN	1930	1972	EGF-LIKE 33, CALCIUM-BINDING.
CC	FIBRILLIN-1 MOLECULES OR WITH OTHER COMPONENTS OF THE	FT	DOMAIN	1973	2012	EGF-LIKE 34, CALCIUM-BINDING.
CC	MICROFIBRILS.	FT	DOMAIN	2013	2054	EGF-LIKE 35, CALCIUM-BINDING.
CC	-/- SIMILARITY: CONTAINS 47 EGF-LIKE DOMAINS.	FT	REPEAT	2055	2126	TGFBBP 6.
CC	-/- SIMILARITY: CONTAINS 7 TGF-BETA BINDING PROTEIN DOMAINS.	FT	DOMAIN	2127	2165	EGF-LIKE 36, CALCIUM-BINDING.
CC	-----	FT	DOMAIN	2166	2205	EGF-LIKE 37, CALCIUM-BINDING.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	FT	DOMAIN	2206	2246	EGF-LIKE 38, CALCIUM-BINDING.
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	FT	DOMAIN	2247	2290	EGF-LIKE 39, CALCIUM-BINDING.
CC	the European Bioinformatics Institute. There are no restrictions on its	FT	DOMAIN	2291	2332	EGF-LIKE 40, CALCIUM-BINDING.
CC	use by non-profit institutions as long as its content is in no way	FT	REPEAT	2333	2401	TGFBBP 7.
CC	modified and this statement is not removed. Usage by and for commercial	FT	DOMAIN	2402	2443	EGF-LIKE 41, CALCIUM-BINDING.
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>	FT	DOMAIN	2444	2484	EGF-LIKE 42, CALCIUM-BINDING.
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	FT	DOMAIN	2485	2523	EGF-LIKE 43, CALCIUM-BINDING.
CC	-----	FT	DOMAIN	2524	2566	EGF-LIKE 44, CALCIUM-BINDING.
DR	EMBL; AF037800; AAD50328.1; .	FT	DOMAIN	2567	2607	EGF-LIKE 45, CALCIUM-BINDING.
DR	HSSP; P35555; IAPJ.	FT	DOMAIN	2607	2647	EGF-LIKE 46, CALCIUM-BINDING.
DR	InterPro; IPR00152; Asx_hydroxyl.	FT	DOMAIN	2648	2687	EGF-LIKE 47, CALCIUM-BINDING.
DR	InterPro; IPR000561; EGF-like.	FT	DISULFID	85	94	BY SIMILARITY.
DR	InterPro; IPR001881; EGF_CA.	FT	DISULFID	89	100	BY SIMILARITY.
DR	InterPro; IPR001438; EGF-II.	FT	DISULFID	102	111	BY SIMILARITY.
DR	InterPro; IPR002212; Fibril-assoc.	FT	DISULFID	119	129	BY SIMILARITY.
DR	Pfam; PF00008; EGF; 45.	FT	DISULFID	123	134	BY SIMILARITY.
DR	Pfam; PF00683; TB; 9.	FT	DISULFID	136	145	BY SIMILARITY.
DR	PRINTS; PR00010; EGFBLD.	FT	DISULFID	150	160	BY SIMILARITY.
DR	SMART; SM00179; EGF_CA; 40.	FT	DISULFID	154	166	BY SIMILARITY.
DR	SMART; SM00001; EGF_like; 6.	FT	DISULFID	168	177	BY SIMILARITY.
DR	PROSITE; PS00010; ASX_HYDROXYL; 41.	FT	DISULFID	250	262	BY SIMILARITY.
DR	PROSITE; PS00022; EGF_1; UNKNOWN_2.	FT	DISULFID	257	271	BY SIMILARITY.
DR	PROSITE; PS01186; EGF_2; 36.	FT	DISULFID	273	286	BY SIMILARITY.
DR	PROSITE; PS01187; EGF_CA; 41.	FT	DISULFID	292	304	BY SIMILARITY.
KW	Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;	FT	DISULFID	299	313	BY SIMILARITY.
KW	Repeat; Signal; Multigene family.	FT	DISULFID	315	328	BY SIMILARITY.
FT	SIGNAL 1 27 POTENTIAL.	FT	DISULFID	453	465	BY SIMILARITY.
FT	CHAIN 28 2871 FIBRILLIN 1.	FT	DISULFID	460	474	BY SIMILARITY.
FT	DOMAIN 81 112 EGF-LIKE 1.	FT	DISULFID	476	488	BY SIMILARITY.
FT	DOMAIN 115 146 EGF-LIKE 2.	FT	DISULFID	494	504	BY SIMILARITY.
FT	DOMAIN 147 178 EGF-LIKE 3.	FT	DISULFID	499	513	BY SIMILARITY.
FT	DOMAIN 246 287 EGF-LIKE 4, CALCIUM-BINDING.	FT	DISULFID	515	528	BY SIMILARITY.
FT	DOMAIN 288 329 EGF-LIKE 5, CALCIUM-BINDING.	FT	DISULFID	534	546	BY SIMILARITY.
FT	REPEAT 330 390 TGFBBP 1.	FT	DISULFID	541	555	BY SIMILARITY.
FT	DOMAIN 392 446 PRO-RICH.	FT	DISULFID	557	570	BY SIMILARITY.
FT	DOMAIN 449 489 EGF-LIKE 6.	FT	DISULFID	576	587	BY SIMILARITY.
FT	DOMAIN 490 529 EGF-LIKE 7, CALCIUM-BINDING.	FT	DISULFID	582	596	BY SIMILARITY.
FT	DOMAIN 530 571 EGF-LIKE 8, CALCIUM-BINDING.	FT	DISULFID	598	611	BY SIMILARITY.
FT	DOMAIN 572 612 EGF-LIKE 9, CALCIUM-BINDING.	FT	DISULFID	617	628	BY SIMILARITY.
FT	DOMAIN 613 653 EGF-LIKE 10, CALCIUM-BINDING.	FT	DISULFID	623	637	BY SIMILARITY.
FT	REPEAT 654 722 TGFBBP 2.	FT	DISULFID	639	652	BY SIMILARITY.
FT	DOMAIN 723 764 EGF-LIKE 11, CALCIUM-BINDING.	FT	DISULFID	727	739	BY SIMILARITY.
FT	DOMAIN 765 806 EGF-LIKE 12, CALCIUM-BINDING.	FT	DISULFID	734	748	BY SIMILARITY.
FT	DOMAIN 807 846 EGF-LIKE 13, CALCIUM-BINDING.	FT	DISULFID	750	763	BY SIMILARITY.
FT	DOMAIN 910 951 EGF-LIKE 14, CALCIUM-BINDING.	FT	DISULFID	769	781	BY SIMILARITY.
FT	REPEAT 952 1027 TGFBBP 3.	FT	DISULFID	776	790	BY SIMILARITY.
FT		FT	DISULFID	792	805	BY SIMILARITY.



the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR	EMBL; U03272; AAA18950.1; -.
DR	EMBL; X62009; -, NOT_ANNOTATED_CDS.
DR	PIR; S17063; S17063.
DR	PIR; S31101; S31101.
DR	HSSP; P35555; IEMN.
DR	Genew; HGNC:3604; FBN2.
DR	MIM; 121050; -.
DR	InterPro; IPR001152; Asx_hydroxyl.
DR	InterPro; IPR000561; EGF-like.
DR	InterPro; IPR001881; EGF-Ca.
DR	InterPro; IPR001438; EGF-II.
DR	Pfam; PF00088; EGF_45.
DR	Pfam; PF00683; TB; 9.
DR	PRINTS; PR0010; EGFBL0OD.
DR	SMART; SM00179; EGF_CA; 43.
DR	SMART; SM00001; EGF_like; 3.
DR	PROSITE; PS00010; ASX_HYDROXYL; 43.
DR	PROSITE; PS00022; EGF_1; 2.
DR	PROSITE; PS01186; EGF_2; 37.
DR	PROSITE; PS01187; EGF_CA; 42.
KW	Extracellular matrix; Calcium-binding; Glycoprotein; EGF-like domain;
KW	Signal; Multigene family; Disease mutation; Polymorphism.
FT	SIGNAL 1 28
FT	CHAIN 29 2911
FT	DOMAIN 111 142
FT	DOMAIN 145 176
FT	DOMAIN 176 207
FT	DOMAIN 275 316
FT	DOMAIN 317 358
FT	REPEAT 359 425
FT	DOMAIN 493 533
FT	DOMAIN 534 573
FT	DOMAIN 574 615
FT	DOMAIN 616 656
FT	DOMAIN 657 697
FT	REPEAT 698 766
FT	DOMAIN 767 808
FT	DOMAIN 809 850
FT	DOMAIN 851 890
FT	DOMAIN 954 995
FT	REPEAT 996 1071
FT	DOMAIN 1072 1113
FT	DOMAIN 1114 1156
FT	DOMAIN 1157 1198
FT	DOMAIN 1199 1240
FT	DOMAIN 1241 1281
FT	DOMAIN 1282 1323
FT	DOMAIN 1324 1365
FT	DOMAIN 1366 1406
FT	DOMAIN 1407 1447
FT	DOMAIN 1448 1489
FT	DOMAIN 1490 1530
FT	DOMAIN 1531 1571
FT	REPEAT 1572 1648
FT	DOMAIN 1649 1690
FT	DOMAIN 1691 1732
FT	REPEAT 1733 1806
FT	DOMAIN 1807 1848
FT	DOMAIN 1849 1890
FT	DOMAIN 1891 1932
FT	DOMAIN 1933 1971
FT	DOMAIN 1972 2014
FT	DOMAIN 2015 2054
FT	DOMAIN 2055 2096
FT	REPEAT 2097 2169
FT	DOMAIN 2170 2211
FT	DOMAIN 2212 2253

FT	DOMAIN	2212	2251	EGF-LIKE 37, CALCIUM-BINDING.
FT	DOMAIN	2292	2292	EGF-LIKE 38, CALCIUM-BINDING.
FT	DOMAIN	2293	2336	EGF-LIKE 39, CALCIUM-BINDING.
FT	DOMAIN	2329	2378	EGF-LIKE 40, CALCIUM-BINDING.
FT	REPEAT	2379	2447	TGFBR 7.
FT	DOMAIN	2448	2489	EGF-LIKE 41, CALCIUM-BINDING.
FT	DOMAIN	2490	2530	EGF-LIKE 42, CALCIUM-BINDING.
FT	DOMAIN	2531	2569	EGF-LIKE 43, CALCIUM-BINDING.
FT	DOMAIN	2570	2612	EGF-LIKE 44, CALCIUM-BINDING.
FT	DOMAIN	2613	2652	EGF-LIKE 45, CALCIUM-BINDING.
FT	DOMAIN	2653	2693	EGF-LIKE 46, CALCIUM-BINDING.
FT	DOMAIN	2694	2733	EGF-LIKE 47, CALCIUM-BINDING.
FT	DISULFID	115	124	BY SIMILARITY.
FT	DISULFID	119	130	BY SIMILARITY.
FT	DISULFID	132	141	BY SIMILARITY.
FT	DISULFID	149	159	BY SIMILARITY.
FT	DISULFID	153	164	BY SIMILARITY.
FT	DISULFID	166	175	BY SIMILARITY.
FT	DISULFID	180	190	BY SIMILARITY.
FT	DISULFID	184	195	BY SIMILARITY.
FT	DISULFID	197	206	BY SIMILARITY.
FT	DISULFID	279	291	BY SIMILARITY.
FT	DISULFID	286	300	BY SIMILARITY.
FT	DISULFID	302	315	BY SIMILARITY.
FT	DISULFID	321	333	BY SIMILARITY.
FT	DISULFID	328	342	BY SIMILARITY.
FT	DISULFID	344	357	BY SIMILARITY.
FT	DISULFID	497	509	BY SIMILARITY.
FT	DISULFID	504	518	BY SIMILARITY.
FT	DISULFID	520	532	BY SIMILARITY.
FT	DISULFID	538	548	BY SIMILARITY.
FT	DISULFID	543	557	BY SIMILARITY.
FT	DISULFID	559	572	BY SIMILARITY.
FT	DISULFID	578	590	BY SIMILARITY.
FT	DISULFID	585	599	BY SIMILARITY.
FT	DISULFID	601	614	BY SIMILARITY.
FT	DISULFID	620	631	BY SIMILARITY.
FT	DISULFID	626	640	BY SIMILARITY.
FT	DISULFID	642	655	BY SIMILARITY.
FT	DISULFID	661	672	BY SIMILARITY.
FT	DISULFID	667	681	BY SIMILARITY.
FT	DISULFID	683	696	BY SIMILARITY.
FT	DISULFID	771	783	BY SIMILARITY.
FT	DISULFID	778	792	BY SIMILARITY.
FT	DISULFID	794	807	BY SIMILARITY.
FT	DISULFID	813	825	BY SIMILARITY.
FT	DISULFID	820	834	BY SIMILARITY.
FT	DISULFID	836	849	BY SIMILARITY.
FT	DISULFID	855	865	BY SIMILARITY.
FT	DISULFID	860	874	BY SIMILARITY.
FT	DISULFID	876	889	BY SIMILARITY.
FT	DISULFID	958	970	BY SIMILARITY.
FT	DISULFID	965	979	BY SIMILARITY.
FT	DISULFID	1076	1088	BY SIMILARITY.
FT	DISULFID	1083	1097	BY SIMILARITY.
FT	DISULFID	1099	1112	BY SIMILARITY.
FT	DISULFID	1118	1130	BY SIMILARITY.
FT	DISULFID	1125	1139	BY SIMILARITY.
FT	DISULFID	1141	1155	BY SIMILARITY.
FT	DISULFID	1161	1173	BY SIMILARITY.
FT	DISULFID	1168	1182	BY SIMILARITY.
FT	DISULFID	1184	1197	BY SIMILARITY.
FT	DISULFID	1203	1215	BY SIMILARITY.
FT	DISULFID	1210	1224	BY SIMILARITY.

Query Match 12.6%; Score 366; DB 1; Length 2911;  
Best Local Similarity 33.1%; Pred. No. 1.4e-16;

```

QY 244 DCSVENGGCEHACNAIPGAPRCQPAGAAIQADGRSCTASATQSCNDLCEHFCVNPDP-- 301
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1244 ECMINNGCDTQCTNSEGSYECSCSEGYALMPDGRSCA-----DIDE--CENNPDIC 1299

```

QY 302 -----OPGSYSCHMCTGTYHLAADQHRCEVDVDCILEPSPCP-QRCVNTQGGFECYCYP 353  
 DB 1294 DGGQCTNIPGTYRCLCYDGFMSDMKTCIDVNECDLNSNCFMGECENTKGSFICHQCL 1353  
 QY 354 NYDLVDGE-CVEPVDC--FRANCEVQCQPLN-QTSYLCVCAEGFAPIPH-----EPHR 403  
 DB 1354 GYSVRKKGTTGCTD-VDECEIGAHCNDWHASCLNIPGSKSCREG--IGNGIKCIDLDE 1410  
 QY 404 COMFCNQTACPADPNTQAS--CECPGTYILDDGFTCTDIDE-----CENG----- 448  
 DB 1411 CSNGTHQCSINACQ-VNTPGYSYACSEGP-TGDFFTCSVDDECAENINLNCNGCQLNVP 1468  
 QY 449 -----GF-----CSGVCHNLPGTFECICGPDSSALARHIG- 477  
 DB 1469 GAYRCECEMGFTPASDRSCQDIDECFSQNICVSGTCNNLPGMFHCICDDGYELDRGTGN 1528  
 QY 478 -TDGD 481  
 DB 1529 CTDID 1533

RESULT 12  
 FBN1\_HUMAN  
 ID FBN1\_HUMAN STANDARD; PRT; 2871 AA.  
 AC P35555;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Fibrillin 1 precursor.  
 GN FBN1 OR FBN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RA Pereira L.V., D'Alessio M., Ramirez F., Lynch J.R., Sykes B.,  
 RA Pangillan T., Bonadio J.;  
 RT "Genomic organization of the sequence coding for fibrillin, the  
 RT defective gene product in Marfan syndrome.";  
 RL Hum. Mol. Genet. 2:961-968(1993).  
 RN [2]  
 RP SEQUENCE OF 1-932 FROM N.A.  
 RC TISSUE=Placenta, and Fibroblast;  
 RA Corson G.M., Chabberg S.C., Dietz H.C., Charbonneau N.L., Sakai L.Y.;  
 RT "Fibrillin binds calcium and is coded by cDNAs that reveal a  
 RT multidomain structure and alternatively spliced exons at the 5'  
 RT end.";  
 RL Genomics 17:476-484(1993).  
 RN [3]  
 RP SEQUENCE OF 899-2871 FROM N.A.  
 RA Maslen C.L., Corson G.M., Maddox B.K., Glanville R.W., Sakai L.Y.;  
 RT "Partial sequence of a candidate gene for the Marfan syndrome.";  
 RL Nature 352:334-337(1991).  
 RN [4]  
 RP SEQUENCE OF 813-1313 FROM N.A.  
 RA Lee B., Godfrey M., Vitale E., Hori H., Mattel M.-G., Sarfarazi M.,  
 RA Tsipouras P., Ramirez F., Hollister D.W.;  
 RT "Linkage of Marfan syndrome and a phenotypically related disorder to  
 RT two different fibrillin genes.";  
 RL Nature 352:330-334(1991).  
 RN [5]  
 RP CHARACTERIZATION.  
 RA Sakai L.Y., Keene D.R., Glanville R.W., Bachinger H.P.;  
 RT "Purification and partial characterization of fibrillin, a cysteine-  
 RT rich structural component of connective tissue microfibrils.";

RL J. Biol. Chem. 266:14763-14770(1991).  
 RN [6]  
 RP STRUCTURE BY NMR OF 2054-2125.  
 RX MEDLINE-98031893; PubMed-9362480;  
 RA Yuan X., Downing A.K., Knott V., Handford P.A.;  
 RT "Solution structure of the transforming growth factor beta-binding  
 RT protein-like module, a domain associated with matrix fibrils.";  
 RL EMBO J. 16:6659-6666(1997).  
 RN [7]  
 RP STRUCTURE BY NMR OF 2124-2205.  
 RX MEDLINE-96144829; PubMed-8568869;  
 RA Knott V., Downing A.K., Cardy C.M., Handford P.A.;  
 RT "Calcium binding properties of an epidermal growth factor-like domain  
 RT pair from human fibrillin-1.";  
 RL J. Mol. Biol. 255:22-27(1996).  
 RN [8]  
 RP STRUCTURE BY NMR OF 2124-2205.  
 RX MEDLINE-96222301; PubMed-8653794;  
 RA Downing A.K., Knott V., Werner J.M., Cardy C.M., Campbell I.D.,  
 RA Handford P.A.;  
 RT "Solution structure of a pair of calcium-binding epidermal growth  
 RT factor-like domains: implications for the Marfan syndrome and other  
 RT genetic disorders.";  
 RL Cell 85:597-605(1996).  
 RN [9]  
 RP REVIEW ON MFS VARIANTS.  
 RX MEDLINE-96174615; PubMed-8594563;  
 RA Colod G., Beroud C., Soussi T., Junien C., Boileau C.;  
 RT "Software and database for the analysis of mutations in the human  
 RT FBN1 gene.";  
 RL Nucleic Acids Res. 24:137-141(1996).  
 RN [10]  
 RP REVIEW ON MFS VARIANTS.  
 RX MEDLINE-97169383; PubMed-9016526;  
 RA Colod-Beroud G., Beroud C., Ades L., Black C., Boxer M., Brock D.J.,  
 RA Godfrey M., Hayward C., Karttunen L., Milewicz D., Peltonen L.,  
 RA Richards R.I., Wang W., Junien C., Boileau C.;  
 RT "Marfan Database (second edition): software and database for the  
 RT analysis of mutations in the human FBN1 gene.";  
 RL Nucleic Acids Res. 25:147-150(1997).  
 RN [11]  
 RP REVIEW ON VARIANTS.  
 RX MEDLINE-98062175; PubMed-9401003;  
 RA Hayward C., Brock D.J.H.;  
 RT "Fibrillin-1 mutations in Marfan syndrome and other type-1  
 RT fibrillinopathies.";  
 RL Hum. Mutat. 10:415-423(1997).  
 RN [12]  
 RP VARIANT MFS PRO-1137.  
 RX MEDLINE-91304569; PubMed-1852208;  
 RA Dietz H.C., Cutting G.R., Pyeritz R.E., Maslen C.L., Sakai L.Y.,  
 RA Corson G.M., Puffenberger E.G., Hamosh A., Nanthakumar E.J.,  
 RA Currstin S.M., Stetten G., Meyers D.A., Franccomano C.A.;  
 RT "Marfan syndrome caused by a recurrent de novo missense mutation in  
 RT the fibrillin gene.";  
 RL Nature 352:337-339(1991).  
 RN [13]  
 RP VARIANTS MFS SER-1249; ARG-1663; SER-2221 AND SER-2307.  
 RX MEDLINE-93250834; PubMed-1301946;  
 RA Dietz H.C., Saraiva J.M., Pyeritz R.E., Cutting G.R., Franccomano C.A.;  
 RT "Clustering of fibrillin (FBN1) missense mutations in Marfan syndrome  
 RT patients at cysteine residues in EGF-like domains.";  
 RL Hum. Mutat. 1:366-374(1992).  
 RN [14]  
 RP VARIANT MFS SER-2307.  
 RX MEDLINE-92235290; PubMed-1569206;  
 RA Dietz H.C., Pyeritz R.E., Puffenberger E.G., Kendzior R.J. Jr.,  
 RA Corson G.M., Maslen C.L., Sakai L.Y., Franccomano C.A., Cutting G.R.;  
 RT "Marfan phenotype variability in a family segregating a missense  
 RT mutation in the epidermal growth factor-like motif of the fibrillin  
 RT gene.";  
 RL J. Clin. Invest. 89:1674-1680(1992).  
 RN [15]

VARIANTS MFS ILE-548 AND ALA-723.  
MEDLINE=94010946; PubMed=8406497;  
Dietz H.C., McIntosh L., Sakai L.Y., Corson G.M., Chalberg S.C.,  
Pyeritz R.E., Francomano C.A.;  
"Four novel FBNI mutations: significance for mutant transcript level  
and EGF-like domain calcium binding in the pathogenesis of Marfan  
syndrome.";  
Genomics 17:468-475(1993).  
[16]  
VARIANTS MFS SER-2144.  
MEDLINE=93278402; PubMed=8504310;  
Hewett D.R., Lynch J.R., Smith R., Sykes B.C.;  
"A novel fibrillin mutation in the Marfan syndrome which could  
disrupt calcium binding of the epidermal growth factor-like module.";  
Hum. Mol. Genet. 2:475-477(1993).  
[17]  
VARIANTS MFS R-862; Y-1117; P-1137 AND F-1589, AND VARIANT A-1148.  
MEDLINE=940108431; PubMed=8281141;  
Tynan K., Comeau K., Pearson M., Wilgenbus P., Levitt D., Gasner C.,  
Berg M.A., Miller D.C., Francke U.;  
"Mutation screening of complete fibrillin-1 coding sequence: report  
of five new mutations, including two in 8-cysteine domains";  
Hum. Mol. Genet. 2:1813-1821(1993).  
[18]  
VARIANTS MFS GLY-217 AND ARG-2627.  
MEDLINE=95067970; PubMed=7977366;  
Karttunen L., Raghunath M., Loenqvist L., Peltonen L.;  
"A compound-heterozygous Marfan patient: two defective fibrillin  
alleles result in a lethal phenotype.";  
Am. J. Hum. Genet. 55:1083-1091(1994).  
[19]  
VARIANT EL LYS-2447.  
MEDLINE=94245249; PubMed=8188302;  
Loenqvist L., Child A., Kainulainen K., Davidson R., Puhakka L.,  
Peltonen L.;  
"A novel mutation of the fibrillin gene causing ectopia lentis.";  
Genomics 19:573-576(1994).  
[20]  
VARIANT MFS CYS-627.  
MEDLINE=94272487; PubMed=8004112;  
Hayward C., Rae A.L., Porteous M.E.M., Logie L.J., Brock L.J.;  
"Two novel mutations and a neutral polymorphism in EGF-like domains  
of the fibrillin gene (FBNI): SSCP screening of exons 15-21 in Marfan  
syndrome patients.";  
Hum. Mol. Genet. 3:373-375(1994).  
[21]  
VARIANT MFS CYS-122.  
MEDLINE=94314977; PubMed=8040326;  
Stahl-Hallengren C., Ukkonen T., Kainulainen K., Kristofersson U.,  
Saxne T., Torqvist K., Peltonen L.;  
"An extra cysteine in one of the non-calcium-binding epidermal growth  
factor-like motifs of the FBNI polypeptide is connected to a novel  
variant of Marfan syndrome.";  
J. Clin. Invest. 94:709-713(1994).  
[22]  
VARIANT MFS TYR-1223.  
MEDLINE=94351682; PubMed=8071963;  
Hewett D.R., Lynch J.R., Child A., Sykes B.C.;  
"A new missense mutation of fibrillin in a patient with Marfan  
syndrome.";  
J. Med. Genet. 31:338-339(1994).  
[23]  
VARIANT MFS HIS-1170.  
MEDLINE=95174777; PubMed=7870075;  
Hayward C., Porteous M.E.M., Brock D.J.H.;  
"A novel mutation in the fibrillin gene (FBNI) in familial  
arachnodactyly.";  
Mol. Cell. Probes 8:325-327(1994).  
[24]  
VARIANTS MFS G-217; N-1023; R-1074; Y-1242; R-1513; E-2127; W-2151;  
K-2447 AND R-2511.  
Kainulainen K., Karttunen L., Puhakka L., Sakai L., Peltonen L.;





```
FT DISULFID 1367 1380 BY SIMILARITY.
FT DISULFID 1374 1389 BY SIMILARITY.
FT DISULFID 1391 1402 BY SIMILARITY.
FT DISULFID 1408 1420 BY SIMILARITY.
FT DISULFID 1415 1429 BY SIMILARITY.
FT DISULFID 1431 1444 BY SIMILARITY.
FT DISULFID 1450 1461 BY SIMILARITY.
FT DISULFID 1456 1470 BY SIMILARITY.
FT DISULFID 1472 1485 BY SIMILARITY.
FT DISULFID 1491 1502 BY SIMILARITY.
FT DISULFID 1497 1511 BY SIMILARITY.
FT DISULFID 1513 1526 BY SIMILARITY.
FT DISULFID 1610 1622 BY SIMILARITY.
FT DISULFID 1617 1631 BY SIMILARITY.
FT DISULFID 1633 1646 BY SIMILARITY.
FT DISULFID 1652 1663 BY SIMILARITY.
FT DISULFID 1658 1672 BY SIMILARITY.
FT DISULFID 1674 1687 BY SIMILARITY.

Query Match 12.4%; Score 362; DB 1; Length 2871;
Best Local Similarity 29.4%; Pred. No. 2.5e-16;
Matches 117; Conservative 26; Mismatches 127; Indels 128; Gaps 21;

QY 212 SAAVAPLGL-----QLMCTAPPAGVOGHWAREAPGAWDCSVENGCGCHACNAIPGAPRC 265
DB 1162 SANLCPHRCVNLGKYOCACNPGVHPHDLFCVVIDECSIMNGCGCTFTNSDGSYEC 1221

QY 266 QCPAGAAQLQADGRSCTASATOSCDLCEHFCVNP-----DQGSYSCMCETGYRL 316
DB 1222 SCQPFALMPDQSRCT-----DIDQ--CEDNPICDGGCTNIPGEYRCLCYDGPMA 1271

QY 317 AADQHRCEVDVDCILESPC--PORCVNTQGFECYHCYNVLDVGE--CVPEVDPD--FR 371
DB 1272 SEDMKTCVDVNECDLNPICLSGTCENTKGSFICHCDMGYSKGKGTCTD--INECEIGA 1330

QY 372 ANCEYOCQPLNOT--SYLCVCAEGFA-----PIPEHPRCQMFQCNQACPADCDPN 420
DB 1331 HNCGRHACTNTAGSFCKSCSPGNGIGDIKTDLDCSNGTHMSQH-----ADC-KN 1382

QY 421 TQAS--CCCEGYILDDGFICTIDEC--EN-----GGF----- 450
DB 1383 TMSGYRCLCKDGY--TGDFCTCDLDCSENLGCGNGLNAPGGYRCECDMGFVPSADG 1441

QY 451 -----CS-----GVCNHLPTFCICGPDPSALARHIG-----TDCDSGK 484
DB 1442 KACEDIDECSLPNICVFTCHNLPLGFLRCECEIGYELDRSGGNCNTDVNECLDPTTCTISGN 1501

QY 485 VDGDSGSGEPSPPTPGS-----TLPPPAVGLV 513
DB 1502 CVN-----TPGSYTCDCSPDELPNTRVGCV 1527

RESULT 14
FBL2_MOUSE
ID FBL2_MOUSE STANDARD; PRF; 1221 AA.
AC P37889; Q9WU12;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Fibulin-2 precursor.
GN FBLN2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 27-35.
RC TISSUE=Fibroblast;
RX MEDLINE=94064787; PubMed=8245130;
RA Pan T.-C., Sasaki T., Zhang R.-Z., Faessler R., Timpl R., Chu M.-L.;
RT "Structure and expression of fibulin-2, a novel extracellular matrix
RT protein with multiple EGF-like repeats and consensus motifs for
RT calcium binding.";
```

```
RL J. Cell Biol. 123:1269-1277(1993).
RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=99337686; PubMed=10406956;
RA Graessel S., Sicot F.-X., Gotta S., Chu M.-L.;
RT "Mouse fibulin-2 gene. Complete exon-intron organization and promoter
RT characterization.";
RL Eur. J. Biochem. 263:471-477(1999).
CC -1- FUNCTION: ITS BINDING TO FIBRONECTIN AND SOME OTHER LIGANDS IS
CC CALCULUM DEPENDENT.
CC -1- SUBUNIT: HOMOTRIMER; DISULFIDE-LINKED.
CC -1- SUBCELLULAR LOCATION: Extracellular matrix.
CC -1- ALTERNATIVE PRODUCTS: At least 2 isoforms; 1 (shown here) and
CC 2/EGF3-less; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: COMPONENT OF BOTH BASEMENT MEMBRANES AND OTHER
CC CONNECTIVE TISSUES.
CC -1- SIMILARITY: CONTAINS 3 ANAPHYLATOXIN-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X75285; CAA53040.1; -
CC EMBL; AF135253; AAD34456.1; -
CC EMBL; AF135239; AAD34456.1; JOINED.
CC EMBL; AF135240; AAD34456.1; JOINED.
CC EMBL; AF135241; AAD34456.1; JOINED.
CC EMBL; AF135242; AAD34456.1; JOINED.
CC EMBL; AF135243; AAD34456.1; JOINED.
CC EMBL; AF135244; AAD34456.1; JOINED.
CC EMBL; AF135245; AAD34456.1; JOINED.
CC EMBL; AF135246; AAD34456.1; JOINED.
CC EMBL; AF135247; AAD34456.1; JOINED.
CC EMBL; AF135248; AAD34456.1; JOINED.
CC EMBL; AF135249; AAD34456.1; JOINED.
CC EMBL; AF135250; AAD34456.1; JOINED.
CC EMBL; AF135251; AAD34456.1; JOINED.
CC EMBL; AF135252; AAD34456.1; JOINED.
CC PIR; A49457; A49457.
CC HSP; P00736; IAPQ.
CC MGD; MGI:95488; Fbln2.
CC InterPro; IPR000020; Anaphylatoxin.
CC InterPro; IPR000152; Asx_hydroxyl.
CC InterPro; IPR000561; EGF-like.
CC InterPro; IPR001881; EGF_Ca.
CC Pfam; PF00008; EGF; 6.
CC SMART; SM00104; ANATO; 2.
CC SMART; SM00179; EGF_CA; 9.
CC SMART; SM00001; EGF_Like; 2.
CC PROSITE; PS00010; ASX_HYDROXYL; 5.
CC PROSITE; PS01177; ANAPHYLATOXIN_1; 3.
CC PROSITE; PS01178; ANAPHYLATOXIN_2; 3.
CC PROSITE; PS00022; EGF_1; FALSE_NEG.
CC PROSITE; PS01186; EGF_2; 5.
CC PROSITE; PS01187; EGF_CA; 10.
KW Signal; Glycoprotein; Extracellular matrix; Plasma; EGF-like domain;
KW Calcium-binding; Alternative splicing; Repeat.
FT SIGNAL 1 26
FT CHAIN 27 1221 FIBULIN-2.
FT DOMAIN 27 434 N.
FT DOMAIN 27 176 SUBDOMAIN NA (CYS-RICH).
FT DOMAIN 177 434 SUBDOMAIN NB (CYS-FREE).
FT DOMAIN 435 477 ANAPHYLATOXIN-LIKE 1.
FT DOMAIN 478 510 ANAPHYLATOXIN-LIKE 2.
FT DOMAIN 511 543 ANAPHYLATOXIN-LIKE 3.
FT DOMAIN 594 635 EGF-LIKE 1, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 669 708 EGF-LIKE 2.
```

```

FT DOMAIN 709 755 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 756 800 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 801 846 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 847 894 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 895 937 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 938 979 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 980 1018 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1019 1061 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1062 1106 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 1111 1221 DOMAIN III.
FT SITE 421 423 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 435 462 BY SIMILARITY.
FT DISULFID 436 469 BY SIMILARITY.
FT DISULFID 449 470 BY SIMILARITY.
FT DISULFID 479 508 BY SIMILARITY.
FT DISULFID 492 509 BY SIMILARITY.
FT DISULFID 511 535 BY SIMILARITY.
FT DISULFID 512 543 BY SIMILARITY.
FT DISULFID 525 542 BY SIMILARITY.
FT DISULFID 598 610 BY SIMILARITY.
FT DISULFID 606 619 BY SIMILARITY.
FT DISULFID 621 634 BY SIMILARITY.
FT DISULFID 673 683 BY SIMILARITY.
FT DISULFID 679 692 BY SIMILARITY.
FT DISULFID 694 707 BY SIMILARITY.
FT DISULFID 713 726 BY SIMILARITY.
FT DISULFID 720 735 BY SIMILARITY.
FT DISULFID 742 754 BY SIMILARITY.
FT DISULFID 805 818 BY SIMILARITY.
FT DISULFID 812 827 BY SIMILARITY.
FT DISULFID 833 845 BY SIMILARITY.
FT DISULFID 899 912 BY SIMILARITY.
FT DISULFID 906 921 BY SIMILARITY.
FT DISULFID 923 936 BY SIMILARITY.
FT DISULFID 942 954 BY SIMILARITY.
FT DISULFID 950 963 BY SIMILARITY.
FT DISULFID 965 978 BY SIMILARITY.
FT DISULFID 984 993 BY SIMILARITY.
FT DISULFID 989 1002 BY SIMILARITY.
FT DISULFID 1004 1017 BY SIMILARITY.
FT DISULFID 1023 1035 BY SIMILARITY.
FT DISULFID 1031 1044 BY SIMILARITY.
FT DISULFID 1046 1060 BY SIMILARITY.
FT DISULFID 1066 1079 BY SIMILARITY.
FT DISULFID 1073 1088 BY SIMILARITY.
FT DISULFID 1093 1105 BY SIMILARITY.
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 497 497 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 737 737 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1072 1072 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 709 755 MISSING (IN ISOFORM 2).
FT CONFLICT 140 159 HSGRKYAGHTVHLSSCRAC -> TVAVSICWPYRPLPLIP
FT CONFLICT 348 348 S -> L (IN REF. 2).
FT CONFLICT 507 507 Q -> QO (IN REF. 2).
FT CONFLICT 1102 1102 Q -> E (IN REF. 2).
SQ SEQUENCE 1221 AA; 131818 MW; 87DB2A10A8FDC45F CRC64;

Query Match
Best Local Similarity 29.5%; Pred. No. 2.4e-16;
Matches 123; Conservative 39; Mismatches 141; Indels 114; Gaps 27;

Qy 154 EQQCEVKAD-GFLCEFHFPATC-----RPLAV-----EPGAAARAVS-----IT 192
Db 521 EGQSCSNPLGYPCN-HVLMSCGEELVPEVRPPPEAAPPVSEMEMASREALS 579
Qy 193 YCT---PFAARGADFO---ALP-----VGSSAAVPLGLQLM-----CTAPPG 229
Db 580 LGTEALPNSLPQDDQDECLMLPGLCOHLICINTVSYRCACPFGLQGDGRTCRPDRG 639
Qy 230 AVQGHWARE-APGAWDCSV-----ENGCEHACNAIPGAPRQCQCPAGAA 272
Db 640 APQLDTARESPRESAQVSPNTIPLVPQPNCKDNGPCRCVGVGDTAMCFCPGA 699

```





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:22:32 ; Search time 31.5 Seconds  
(without alignments)  
3375.248 Million cell updates/sec

Title: US-09-509-994-2

Perfect score: 2916  
Sequence: 1 MLGVLVIGALALAGLGFAP.....PSPTPGSTLTPPAVLVHSG 516

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_21.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	2658	91.2	468	4 Q9UC32	Q9uc32 homo sapien
2	1844	63.2	577	11 Q35370	Q35370 rattus norv
3	1400.5	48.0	461	11 P97883	P97883 rattus norv
4	584.5	20.0	757	4 Q9HC00	Q9hcu0 homo sapien
5	575	19.7	765	11 Q91ZV1	Q91zv1 mus musculu
6	575	19.7	765	11 Q91V98	Q91v98 mus musculu
7	431	14.8	1664	5 Q9TVQ2	Q9tvg2 caenorhabdi
8	414	14.2	1574	11 Q88281	Q88281 rattus norv
9	386	13.2	1394	5 Q9VS89	Q9vs89 drosophila
10	376	12.9	708	13 P87363	P87363 gallus gall
11	375	12.9	2906	11 Q9WU99	Q9wuh9 rattus norv
12	369	12.7	1246	4 Q75095	Q75095 homo sapien
13	367	12.6	3857	11 Q88840	Q88840 mus musculu
14	365	12.5	2809	4 Q96JP8	Q96jp8 homo sapien
15	361	12.4	2872	11 Q9WU98	Q9wuh8 rattus norv
16	359	12.3	528	11 Q9CXD8	Q9cxd8 mus musculu

17	358.5	12.3	1174	11 Q99K58	Q99k58 mus musculu
18	357	12.2	937	4 Q96FT5	Q96ft5 homo sapien
19	353.5	12.1	746	4 Q96HB9	Q96hb9 homo sapien
20	353.5	12.1	1256	4 Q9NS15	Q9nsl5 homo sapien
21	353.5	12.1	1382	4 Q9H7K2	Q9h7k2 homo sapien
22	352.5	12.1	741	4 Q96R89	Q96r89 homo sapien
23	351.5	12.1	1511	4 Q75412	Q75412 homo sapien
24	351.5	12.1	1587	4 Q00508	Q00508 homo sapien
25	342	11.7	5636	4 Q96RW7	Q96rw7 homo sapien
26	341	11.7	576	4 Q9Y3V7	Q9y3v7 homo sapien
27	340	11.7	1062	11 Q60789	Q60789 mus musculu
28	338.5	11.6	1289	5 Q8SSS3	Q8sss3 dictyosteli
29	336	11.5	2673	4 Q96SC3	Q96sc3 homo sapien
30	333.5	11.4	961	11 Q9EQC6	Q9eqc6 mus musculu
31	332.5	11.4	1764	11 Q35806	Q35806 rattus norv
32	332	11.4	999	4 Q9NQ36	Q9nq36 homo sapien
33	331.5	11.4	1713	11 Q88349	Q88349 mus musculu
34	331	11.4	589	5 Q9T2S1	Q9t2s1 caenorhabdi
35	329.5	11.3	956	11 Q8R542	Q8r542 mus musculu
36	329	11.3	2189	5 Q9BI05	Q9bi05 eimeria ten
37	328.5	11.3	956	11 Q99K64	Q99k64 mus musculu
38	325	11.1	1095	11 Q60784	Q60784 mus musculu
39	324.5	11.1	937	5 Q9BLJ1	Q9blj1 ciona intes
40	323	11.1	1253	11 Q61810	Q61810 mus musculu
41	322.5	11.1	1394	4 Q8TD95	Q8td95 homo sapien
42	319.5	11.0	1833	11 Q08999	Q08999 mus musculu
43	315.5	10.8	443	4 Q96TF5	Q96tf5 homo sapien
44	314.5	10.8	3507	5 Q23587	Q23587 caenorhabdi
45	314	10.8	997	11 Q9JJS0	Q9jjs0 mus musculu

ALIGNMENTS

RESULT 1

Q9UC32	PRELIMINARY;	PRT;	468 AA.
ID	Q9UC32		
AC	Q9UC32;		
DT	01-MAY-2000 (Tremblrel. 13, Created)		
DT	01-MAY-2000 (Tremblrel. 13, Last sequence update)		
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)		
DE	THROMBOMODULIN.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE.		
RX	MEDLINE=93293792; PubMed=8390446;		
RA	Yamamoto S., Mizoguchi T., Tamaki T., Ohkuchi M., Kimura S., Aoki N.;		
RT	"Urinary thrombomodulin, its isolation and characterization.;"		
RL	J. Biochem. 113:433-440(1993).		
DR	HSP; P07204; IZAQ.		
DR	InterPro; IPR000152; Asx_hydroxyl.		
DR	InterPro; IPR000561; EGF-like.		
DR	InterPro; IPR001881; EGF_Ca.		
DR	InterPro; IPR001304; Lectin_C.		
DR	InterPro; IPR001491; Thrbomoduln.		
DR	Pfam; PF00008; EGF_4.		
DR	Pfam; PF00059; Lectin_c; 1.		
DR	PRINTS; PR00907; THRBOMODULN.		
DR	SMART; SM00034; CLEFT; 1.		
DR	SMART; SM00181; EGF; 6.		
DR	PROSITE; PS00010; ASX_HYDROXYL; 1.		
DR	PROSITE; PS00041; C-TYPE-LECTIN_2; 1.		
DR	PROSITE; PS01186; EGF_2; 2.		
DR	PROSITE; PS01187; EGF_CA; 1.		
KW	Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.		
SQ	SEQUENCE 468 AA; 49444 MW; 4BF8E98FEF86A40 CRC64;		

Query Match 91.2%; Score 2658; DB 4; Length 468;  
Best Local Similarity 99.8%; Pred. NO. 8.4e-205;  
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 19 APAEPQGGSCVCEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 78
DB 1 APAEPQGGSCVCEHDCFALYPGPATFLNASQICDGLRGHLMTVRSSVAADVLSLLNGD 60
QY 79 GVGRRRLWIGLQPLPGCGDPKRLGRLGFWGTGDNNTSYSRWARLDLNGAPLCPGLCV 138
DB 61 GVGRRRLWIGLQPLPGCGDPKRLGRLGFWGTGDNNTSYSRWARLDLNGAPLCPGLCV 120
QY 139 AVSAAEATVPSEPTWEEQCEVADGFLCEHFFPFCRPLAVEPFGAAAAVSYTYGTPFA 198
DB 121 AVSAAEATVPSEPTWEEQCEVADGFLCEHFFPFCRPLAVEPFGAAAAVSYTYGTPFA 180
QY 199 ARGADFQALPVGSSAAVAPLGLQMLCTAPGAVQGHAREAPGAWDCSVENGCEHACNA 258
DB 181 ARGADFQALPVGSSAAVAPLGLQMLCTAPGAVQGHAREAPGAWDCSVENGCEHACNA 240
QY 259 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 318
DB 241 IPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNPDPQPGSYSCMCETGYRLAA 300
QY 319 DQHRCEVDVDCILEPSPCQRCVNTQGGFCHCYPNVDLVGCEVPEVDFCFRANCEYQC 378
DB 301 DQHRCEVDVDCILEPSPCQRCVNTQGGFCHCYPNVDLVGCEVPEVDFCFRANCEYQC 360
QY 379 QPLNQTSLVLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPNTQASCECEGYILDGFI 438
DB 361 QPLNQTSLVLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDPNTQASCECEGYILDGFI 420
QY 439 CTDIDECENGFCGVCNHLNPGTPECICGPDPSALAHIGTDCDSGKYD 486
DB 421 CTDIDECENGFCGVCNHLNPGTPECICGPDPSALAHIGTDCDSGKYD 468

RESULT 2
O35370 PRELIMINARY; PRT; 577 AA.
AC O35370; 1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE THROMBOMODULIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LUNG;
RA Wang J., Yao A., Wang J.Y., Hardin J.W., Hauer-Jensen M.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 1-136 FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Yao A., Wang J., Hardin J.W., Hauer-Jensen M.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF022743; AAB80760.1; -.
DR EMBL; AF022742; AAB80923.1; -.
DR HSSP; P07204; 1FGD.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR001491; Thrombomodulin.
DR Pfam; PF00008; EGF; 5.
DR PRINTS; PR00907; THROMBOMODULN.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00001; EGF_LIKE; 3.
DR PROSITE; PS00010; ASX_HYDROXYL; 2.
DR PROSITE; PS00041; C-TYPE-LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 3.
```

```
DR PROSITE; PS01187; EGF_CA; 2.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 577 AA; 61844 MW; 0BE764CBF18555F CRC64;

Query Match
Best Local Similarity 64.0%; Score 1844; DB 11; Length 577;
Matches 333; Conservative 49; Mismatches 132; Indels 6; Gaps 5;

QY 1 MLGVVLGLALAGLFGPAPAEPPQGGSCVCEHDCFALYPGPATFLNASQICDGLRGHL 60
DB 1 MLGVVLGLVAPLAGLISALAKLPKQSCVGNCECFALFQDPVTELDASQACORLQGHLM 60
QY 61 TYRSSVAADVLSLLNGDGVGRRRLWIGLQPLPGCGDPKRLGRLGFWGTGDNNTSYS 120
DB 61 TYRSSVAADVLSLLVS-DSSMDSRP-WIGLQPLPGCGDPVHLGRLGFWGTGDNNTSYS 118
QY 121 RWARLDLNGAPLCPGLCVAAVSAEATVPSEPTWEEQCEVADGFLCEHFFPFCRPLAV 180
DB 119 RWARNDSQSPPLCGPLCVTVSTATAAPEPAWEEKPCENETKGLCFEYFAAFCRPLRV 178
QY 181 EP-GAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPGAVQGHAREA 239
DB 179 NTRDPEGAHISSTYNTPLGVSAGADFQTLPGSSATVAPFGLVLCRALPGTSEGHWTREV 238
QY 240 PGAWDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 299
DB 239 TGAWNCVSENGCEYCNRSANGPRCPCVPSGGDLQADGRSCKAPVAQJCNELCQHFVNN 298
QY 300 PQPGSYSCMCETGYRLAADHRCEDVDVDCILEPSPCQRCVNTQGGFCHCYPNVDLV 359
DB 299 SDVPGSYSCMCETGYRLAADHRCEDVDVDCILEPSPCQRCVNTQGGFCHCYPNVDLV 358
QY 360 GECVEVDPCFRANCEYQCQPLNQTSLVLCVCAEGFAPIPHEPHRCOMFCNOTACPADCDP 419
DB 359 GECVEQLDPCFRSKCEYQCPVNSTHYNCIAEGFAPKLDLDDPDRCEMFCNETSCPADCDP 418
QY 420 NTQASCECEPEGYILDGFICTDIDECENGFCGVCNHLNPGTPECICGPDPSALAHIGT 479
DB 419 NSPSFCQCEPFGILDEGSICTDIDECQSGECLTNECRNLPGSYECICGPDPTALAGQISK 478
QY 480 CDSGKY--DGGDSGSGEPPPS-PTFGSTLTTPAVGLVHSG 516
DB 479 CDPIPVLESDSGSGEHPSSNPTVTVSVTPSARPMHSG 518

RESULT 3
P97883 PRELIMINARY; PRT; 461 AA.
ID P97883
AC P97883;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Thrombomodulin (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN CAPILLARY;
RA Wang L., Tran N.D., Schreiber S.S., Zlokovic B.V.;
RT "Nucleotide sequence of rat thrombomodulin.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U90121; AAB49723.1; -.
DR HSSP; P07204; 1FGD.
DR InterPro; IPR00152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001491; Thrombomodulin.
DR Pfam; PF00008; EGF; 5.
DR PRINTS; PR00907; THROMBOMODULN.
DR SMART; SM00179; EGF_CA; 2.
DR SMART; SM00001; EGF_LIKE; 3.
```

```
DR PROSITE: PS00010; ASX_HYDROXYL; 2.
DR PROSITE: PS01186; EGF_2; 3.
DR PROSITE: PS01187; EGF_CA; 2.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
FT NON_TER 1
SQ SEQUENCE 461 AA; 49499 MW; 45971BCE84688E67 CRC64;

Query Match 48.0%; Score 1400.5; DB 11; Length 461;
Best Local Similarity 62.7%; Pred. No. 3.9e-104;
Matches 245; Conservative 37; Mismatches 104; Indels 5; Gaps 4;

QY 131 PLCGPLCAVAAETVPSPEIWEQCEVKADGFLCEFFHFPATC-RPLAVEP-GAAAAA 188
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 12 PLCGPLCVTVSTATEAAGEPAWEKPCENETKGFCEFFFAFCRPLRVNTRDEGAH 71
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 189 VSITYGTPTFAARGADQALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAPAGWDCSVE 248
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 72 ISSTYNTPLGVSGADQTLPIGSSATVAPFGLVLCRALPGTSEGHWTREVTGAWNCVSVE 131
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 249 NGGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFVCPNPPQSGSYSC 308
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 132 NGGCEYMCNRSANGPRCVCPGSGDLQADGRSCAKPYGLQCLNELCOHFCVNNSDVPQSGSYSC 191
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 309 MCETGYRLAADHRCEDVDCCILEPSPCPORCVNTGGFECHCYPNVLDLVGCEVPEVDP 368
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 192 MCETGYRLAADHRCEDVDCCILEPSPCPORCVNTGGFECHCYPNVLDLVGCEVPEVDP 251
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 369 FRANCEYQCPUNQTSYLCVCAEGFAPIPHEPHRCQMFQNCQACADCDPNTQASCECP 428
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 252 CFRSKEYQCPVNSYHNCICAEAGFAPKLDLDDRECMFCNETSCPADCDPNSFCQCP 311
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 429 EGYLDGDFICTDIDCENGFGSCGVCHNLPGTFTEICGPDLSALRHIGTDCDSGKY--D 486
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 312 EGYLDGSGICTDIDCSQCELTNECRNLPGSYECICGPDLTALQISKDCDPIPVLED 371
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 487 GDSGSGEPPPS-PTPGSTLTTPPAVGLVHSG 516
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 372 SEDGSGEHPSSNPTVYVSTVPPSARPMHSG 402
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
Q9HCU0 PRELIMINARY; PRT; 757 AA.
AC Q9HCU0;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Tumor endothelial marker 1 precursor (Endosomal protein).
GN TEM1 OR ENDOSIALIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20407466; PubMed=10947988;
RA St Croix B., Rago C., Velculescu V., Traverso G., Romans K.E.,
RA Montgomery E., Lal A., Riggins G.J., Lengauer C., Vogelstein B.,
RA Kinzler K.W.;
RT "Genes expressed in human tumor endothelium.";
RL Science 289:1197-1202(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21269274; PubMed=11084048;
RA Christian S., Ahorn H., Koehler A., Eisenhaber F., Rodi H.P.,
RA Garin-Chesa P., Park J.E., Rettig W.J., Lenter M.C.;
RT "Molecular cloning and characterization of Endosialin, a C-type
RT lectin-like cell surface receptor of tumor Endothelium.";
RL J. Biol. Chem. 276:7408-7414(2001).
DR EMBL; AF279142; AAGO0867.1; -.
DR EMBL; AJ295846; CAC34381.1; -.
DR HSSP; P07204; 1ZAO.
DR InterPro; IPR000152; Asx_hydroxyl.
```

```
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_CA.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00179; EGF_CA; 2.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS01187; EGF_CA; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 757 TUMOR ENDOTHELIAL MARKER 1.
SQ SEQUENCE 757 AA; 80858 MW; C96363EAF8FFA0 CRC64;

Query Match 20.0%; Score 584.5; DB 4; Length 757;
Best Local Similarity 33.6%; Pred. No. 1.4e-38;
Matches 177; Conservative 44; Mismatches 191; Indels 115; Gaps 24;

QY 1 MLGVLTGALALAGLG-FPAPEPQPGSQCVHEHDCFALYPGPATFLNASOICDGLRHL 59
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2 LLRLLLAWAAAGPTLGQDPWAAEPR---AAGPSSCYALFPRRRTFLEAWRACRELGGDL 58
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 60 MTRVSSVAADVVISLLNGDGVGRRRLWIGLQLPFGCDPKRLGPLRGFQVWTDNNITSY 119
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 59 ATPPTPEAQRVDSLVG--AGPASRLWIGLQRAQRCQLQR--PLRGFTWTGDDQDTAF 114
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 120 SWARLDLNGAPICGPLCAVSAAEATVPSPITWEEOQCEVKADGFLCEFFHFPATCRPLA 179
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 115 TNWAQ-PASGGPCPAQRCVALEAS-----GHRWLEGSLTAVDGYLCQFGFEGACPALQ 168
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 180 VEPGAAAAAVSITYGTPTFAARGADQALPVGSSAAV---APLGLQLMCTAPPAGVQGHWA 236
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 169 DEAGQAGPAV--YTPPHLVSTEFELWPFGSVAAVQCAQGRGASLLCVKQPEGGVG-WS 224
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 237 REAP---GAWDCSVENGCEHAC-NAIPGAPRCQCPAGAAQADGRSCTASATQS--CNDL 291
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 225 RAGPLCLGT-GCSPDNGGCEHCEVEVDGHVSCRCFRLAARDGSCDEDCQAQAPCEQ 283
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 292 CEHFCVNPDPQGSYSCMCTGYRLAADQ-HRCEVDVDCILEPSPCPORCVNTQGGFECH 350
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 284 CE-----PGGPGYSCHRLGRFPAEDDPRHCVDTDEQI--AGVCQMCVNVYGGFECY 336
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 351 CYPNYDLVDGCEVPEVPCFRANCEYQCPNLQTSYLCVCAEGFAPIPHEPHRCQMFQNCQ 410
DQ :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 337 CSEGHELE-----ADGIS-----ADGIS-----ADGIS-----ADGIS 349
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 411 TACPADCDP---NTQASCECPGYLDDGFTCTDIDE---CENGFGSCGVCHNLPGT-- 461
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 350 -----CSPAGAMGAQASQDLGD-ELLDDGDEDEDEDEAKAFNGWT-----EMPGLW 397
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 462 FECICGPDLSALRHIGTDCDSKVDGSGSGEPPSPPTPGSTLTTPP 508
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 398 MEPTQPDFALAYRPSFPED-----REFQIPYPEPTWPPP 432
DQ ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
Q912V1 PRELIMINARY; PRT; 765 AA.
AC Q912V1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Endostalin.
GN TEM1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=SWISS WEBSTER; TISSUE=THYMUS;
RX MEDLINE=21486432; PubMed=11489895;
RA Opavsky R., Haviernik P., Jurkovicova D., Garin M.T., Copeland N.G.,
RA Gilbert D.J., Jenkins N.A., Bies J., Garfield S., Pastorekova S.,
RA Oue A., Wolff L.;
RT "Molecular Characterization of the Mouse Tem1/endosomal Gene
RT Regulated by Cell Density in Vitro and Expressed in Normal Tissues in
RT Vivo.";
RL J. Biol. Chem. 276:38795-38807(2001).
RC STRAIN=12956/SVEVITAC; TISSUE=SPLEEN;
RX MEDLINE=21486432; PubMed=11489895;
RA Opavsky R., Haviernik P., Jurkovicova D., Garin M.T., Copeland N.G.,
RA Gilbert D.J., Jenkins N.A., Bies J., Garfield S., Pastorekova S.,
RA Oue A., Wolff L.;
RT "Molecular Characterization of the Mouse Tem1/endosomal Gene
RT Regulated by Cell Density in Vitro and Expressed in Normal Tissues in
RT Vivo.";
RL J. Biol. Chem. 276:38795-38807(2001).
DR EMBL; AF388572; AAK84664.1; -.
DR MGI; MGI:1917695; Tem1.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00084; sushi_1.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS01187; EGF_Ca; UNKNOWN_1.
SQ SEQUENCE 765 AA; 81823 MW; 07B3225E5DD13A03 CRC64;

Query Match 19.7%; Score 575; DB 11; Length 765;
Best Local Similarity 38.3%; Pred. No. 8.4e-38;
Matches 144; Conservative 41; Mismatches 155; Indels 36; Gaps 17;

QY 1 MLGVLVGLALAGLGFPAEPAPGPGSQCVHEHDFALYGPATFLNASQICDGLRGHLM 60
DB 1 LRLLLAWVAAPALG-QVPTPEPRAA-CGPPSSCYALFPRRTFLEAWRACRELGGNLA 59
QY 61 TVRSSVAADVVISLLNGDGGVRRR--LWIGLQPPGCGDKRLGRLGFQWVTGDNNTS 118
DB 60 TPRTPEAQRVDSL--GVGPANGLLWIGLQROAROCQQR--PLRGFIWTTGQDQTA 113
QY 119 YSRWARLDLNGAPLCPGLCVAVSAABATVPSEPIWEEQCEVKADGFLCEFFHPATCRPL 178
DB 114 FTNWAQATEG-PCPAQRCAALEAS-----GEHRWLEGSCTILAVDGYLCQFGFEGACPAL 167
QY 179 AVEPGAAAAVSITYCTPFAARGADFOALPVGSSAAV---APLGLQMTCTAPGAVQGHW 235
DB 168 PLEVQAGPAV---YTPFNLVSSEFWLPGFSAVAVQCGRGLASLLCVKQPSGGVG-W 223
QY 236 AREAP--GAWDCSVENGCGEHAC-NAIPGAPRCQCPAGAAQADGRSCTASATQS-CNDL 291
DB 224 SQTGCLPCTGCGPDNGGCEHECVEDGAVSCRCSEGFRLAADGHSCEDPCAQAPCEQQ 283
QY 292 CEHFCVNPDPQGSYSQCMCTGYRLAADO-HRCEDVDDCILEPSPQRCRVNTQGGFEC 350
DB 284 CE-----PGPGQGSCHCRGLFRPAEDDPHRCVDTDEQI-AGVCQCMCVNYVGGFECY 336
QY 351 CYPNYDL-VDGECVEP 365
DB 337 CSEHGELEADGISCSP 352

RESULT 6
Q91V98 PRELIMINARY; PRT; 765 AA.
AC Q91V98;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Tumor endothelial marker 1 precursor (Endosomal).
GN TEM1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RX MEDLINE=21443268; PubMed=11559528;
RA Carson-Walter E.B., Watkins D.N., Nanda A., Vogelstein B.,
RA Kinzler K.W., St Croix B.;
RT "Cell surface tumor endothelial markers are conserved in mice and
RT humans.";
RL Cancer Res. 61:6649-6655(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=12956/SVEVITAC; TISSUE=SPLEEN;
RX MEDLINE=21486432; PubMed=11489895;
RA Opavsky R., Haviernik P., Jurkovicova D., Garin M.T., Copeland N.G.,
RA Gilbert D.J., Jenkins N.A., Bies J., Garfield S., Pastorekova S.,
RA Oue A., Wolff L.;
RT "Molecular Characterization of the Mouse Tem1/endosomal Gene
RT Regulated by Cell Density in Vitro and Expressed in Normal Tissues in
RT Vivo.";
RL J. Biol. Chem. 276:38795-38807(2001).
DR EMBL; AF388572; AAK84665.1; -.
DR MGI; MGI:1917695; Tem1.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001304; Lectin_C.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF_2.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sushi_1.
DR PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE; PS00041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS01186; EGF_2; UNKNOWN_1.
DR PROSITE; PS01187; EGF_Ca; UNKNOWN_1.
KW Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 765 TUMOR ENDOTHELIAL MARKER 1.
SQ SEQUENCE 765 AA; 81813 MW; 572A06CC15BC8CD8 CRC64;

Query Match 19.7%; Score 575; DB 11; Length 765;
Best Local Similarity 38.3%; Pred. No. 8.4e-38;
Matches 144; Conservative 41; Mismatches 155; Indels 36; Gaps 17;

QY 1 MLGVLVGLALAGLGFPAEPAPGPGSQCVHEHDFALYGPATFLNASQICDGLRGHLM 60
DB 2 LRLLLAWVAAPALG-QVPTPEPRAA-CGPPSSCYALFPRRTFLEAWRACRELGGNLA 59
QY 61 TVRSSVAADVVISLLNGDGGVRRR--LWIGLQPPGCGDKRLGRLGFQWVTGDNNTS 118
DB 60 TPRTPEAQRVDSL--GVGPANGLLWIGLQROAROCQQR--PLRGFIWTTGQDQTA 113
QY 119 YSRWARLDLNGAPLCPGLCVAVSAABATVPSEPIWEEQCEVKADGFLCEFFHPATCRPL 178
DB 114 FTNWAQATEG-PCPAQRCAALEAS-----GEHRWLEGSCTILAVDGYLCQFGFEGACPAL 167
QY 179 AVEPGAAAAVSITYCTPFAARGADFOALPVGSSAAV---APLGLQMTCTAPGAVQGHW 235
DB 168 PLEVQAGPAV---YTPFNLVSSEFWLPGFSAVAVQCGRGLASLLCVKQPSGGVG-W 223
QY 236 AREAP--GAWDCSVENGCGEHAC-NAIPGAPRCQCPAGAAQADGRSCTASATQS-CNDL 291
DB 224 SQTGCLPCTGCGPDNGGCEHECVEDGAVSCRCSEGFRLAADGHSCEDPCAQAPCEQQ 283
QY 292 CEHFCVNPDPQGSYSQCMCTGYRLAADO-HRCEDVDDCILEPSPQRCRVNTQGGFEC 350
DB 284 CE-----PGPGQGSCHCRGLFRPAEDDPHRCVDTDEQI-AGVCQCMCVNYVGGFECY 336
QY 351 CYPNYDL-VDGECVEP 365
DB 337 CSEHGELEADGISCSP 352

RESULT 7
Q9TVQ2 PRELIMINARY; PRT; 1664 AA.
ID Q9TVQ2
```



AC	Q9TV02;	
DT	01-MAY-2000 (TrEMBLrel. 13, Created)	
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)	
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)	
DE	Y64G10A.7 protein.	
GN	Y64G10A.7	
OS	Caenorhabditis elegans.	
OC	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;	
OC	Rhabditidae; Peloderinae; Caenorhabditis.	
OX	NCBI_TaxID=6239;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Mortimore B.J.;	
RL	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=99069613; PubMed=9851916;	
RA	none;	
RT	"Genome sequence of the nematode C.elegans: A platform for	
RT	investigating biology."	
RL	Science 282:2012-2018(1998).	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RA	Ainscough R.;	
RL	Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AL117206; CAB60454.1; -.	
DR	EMBL; AL110498; CAB60454.1; JOINED.	
DR	EMBL; AL110498; CAB57911.1; -.	
DR	EMBL; AL117206; CAB57911.1; JOINED.	
DR	HSSP; P00736; IAPQ.	
DR	InterPro; IPR000152; Asx_hydroxyl.	
DR	InterPro; IPR000561; EGF-like.	
DR	InterPro; IPR001881; EGF_Ca.	
DR	InterPro; IPR002049; Laminin_EGF.	
DR	Pfam; PF00008; EGF; 25.	
DR	PRINTS; PR00011; EGF_LAMININ.	
DR	SMART; SM00179; EGF_Ca; 4.	
DR	SMART; SM00001; EGF_like; 18.	
DR	PROSITE; PS00010; ASX_HYDROXYL; 4.	
DR	PROSITE; PS00022; EGF_1; UNKNOWN_22.	
DR	PROSITE; PS01186; EGF_2; 24.	
DR	PROSITE; PS01187; EGF_Ca; 3.	
KW	Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.	
SQ	SEQUENCE 1664 AA; 179279 MW; A69F093B4C705932 CRC64;	
Query Match 14.8%; Score 431; DB 5; Length 1664;		
Best Local Similarity 29.7%; Pred. No. 6.7e-26;		
Matches 126; Conservative 40; Mismatches 178; Indels 80; Gaps 19;		
QY	119 YSRWRLDNLGAPLCPLCVAVSAE-----ATVPSEPIWEEQCE--VKADGFLCEF 169	
DB	56 YLRFARFSRRGCKCLLRVQANCSADLCHNGGTCTVPSEHNDNEQVCECPVGTGAKQY 115	
QY	170 HPPATCRPLAVEPFAAAAAVSTYGTTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPP 229	
DB	116 D-ANEC--MANNGGCEHCYN-TIGTY-----CRWPG 145	
QY	230 AVQGHWAREAPAGWDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTA----- 282	
DB	146 FELSGDGNCTSDIDCAVSNCGCDRCVNSPGGFRCDPCLYLDHAGRTCGSGFFHENL 205	
QY	283 ----SATQSC---NDLCEHFCVNPDPGYSVCMETGYRLADQHRCEVDVDDCILPSPC 336	
DB	206 ILIKKVTSCSTDNGGCEHC-ENDSNGEYRCRCRVGFKLSENKRSCQPVDPDFDNKGGC 264	
QY	337 PORCVNTQGGFECHCYPNYDL-VGECVEPVPDPCFRAN-CEYQCQPLNQTSLYLCVCAEGF 394	
DB	265 QHCHTNHGRACQCPGPGFHLSDYRRSCVDIDECANNGCEHCENVKGT-YRCKREGY 323	
QY	395 APIPEPHRCQMF-----CNQATCAPDC--DPNTQASCECEGYIL-DDGFICTDIDEC- 445	
DB	324 -QLGRDGRTEEMGLGCQVNGCGCHDCYDQPDGPHVCKRNGYILANDQKLCHDINECH 382	
QY	446 ENGFGCSGVCHNLPGTFECICGPDSSALARIHGT-----DCDS-----GKVDGDSG 491	
DB	383 ENNGDCSQICVNLNLAGSVEQCKPGRFLMKDRKTCEDISECSNNGCGEQICSNQEGGYMC 442	
QY	492 SGEF 495	
DB	443 SCEP 446	
RESULT 8		
QY	088281	
ID	088281 PRELIMINARY; PRT; 1574 AA.	
AC	088281;	
DT	01-NOV-1998 (TrEMBLrel. 08, Created)	
DT	01-NOV-1998 (TrEMBLrel. 08, Last sequence update)	
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)	
DE	MEGF6.	
GN	MEGF6.	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI_TaxID=10116;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	STRAIN=SPRAGUE-DAWLEY; TISSUE=BRAIN;	
RX	MEDLINE=98360089; PubMed=9693030;	
RA	Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;	
RT	"Identification of high-molecular-weight proteins with multiple EGF-	
RT	like motifs by motif-trap screening."	
RL	Genomics 51:27-34(1998).	
DR	EMBL; AB011532; BAA32462.1; -.	
DR	HSSP; P00736; IAPQ.	
DR	InterPro; IPR000152; Asx_hydroxyl.	
DR	InterPro; IPR000561; EGF-like.	
DR	InterPro; IPR001881; EGF_Ca.	
DR	Pfam; PF00008; EGF; 24.	
DR	SMART; SM00179; EGF_Ca; 4.	
DR	SMART; SM00001; EGF_like; 19.	
DR	PROSITE; PS00010; ASX_HYDROXYL; 5.	
DR	PROSITE; PS00022; EGF_1; UNKNOWN_23.	
DR	PROSITE; PS01186; EGF_2; 23.	
DR	PROSITE; PS01187; EGF_Ca; 5.	
KW	Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.	
SQ	SEQUENCE 1574 AA; 165445 MW; 2B48533D8F77F6E7 CRC64;	
Query Match 14.2%; Score 414; DB 11; Length 1574;		
Best Local Similarity 34.1%; Pred. No. 1.4e-24;		
Matches 104; Conservative 28; Mismatches 101; Indels 72; Gaps 17;		
QY	224 CTAPPG-AYQGHWAREAPAGWDCSVENGCEHACNAIPGAPRCQCPAGALQADGRSCTA 282	
DB	147 CRCPGPGYLGQD-GKTCQDVDECRAHNGCGOHCRCVNTPGSYLCECKPGRFLHTDGRCTL- 204	
QY	283 SATQSC---NDLCEHFCVP-----NPDQPGSYSCM---- 309	
DB	205 -ATSSCTLNGSGCHQCHQCVLTVTHRCQCRPOYQLOEDGRRCVRRSPCAENGSGCMHCQ 263	
QY	310 -----CETGYRLAADQHRCEVDVDDCILPSPCQRCVNTQGGFECHCYPNYDL-VD 359	
DB	264 ELRGLAHCGCHPGYQLAADRKTCEDVDICALGLAQCAHGLNTQGSFKVCVCHAGYELGAD 323	
QY	360 G-EC---VFPVDPFCFRAN-CEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMF---- 408	
DB	324 GRCQYRIEMEIVNSCEAGNGCGSHGCSH-TSTGPLCTCPRGY-----ELDEDQKTCIDID 377	
QY	409 ---NOTACPADCPNTQA--SCEPGEYILD-DGFICTDIDECENG-GFCSGVCHNLPGT 461	
DB	378 DCANSPCCQAC-ANTPGYECSCFAGYRLNTDGCCEVDDECASGHGCGEHCNSLAGS 436	
QY	462 FECIC 466	
DB	437 FQFCF 441	

```
RESULT 9
Q9VS89 PRELIMINARY; PRT; 1394 AA.
AC Q9VS89;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG7526 protein (fragment).
GN CG7526.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brattler P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Paulos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissendbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang X., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT The genome sequence of Drosophila melanogaster.;
RL Science 287:2185-2195(2000).
DR HSSL; AE003558; AAF50538.1; -.
DR HSSP; P00736; IAPQ.
DR FlyBase; FBgn0035798; CG7526.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF-Ca.
DR InterPro; IPR003410; Hyalin.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR InterPro; IPR001491; Thrbomodulin.
DR Pfam; PF00008; EGF; 11.
DR Pfam; PF00084; sushi; 2.
DR PRINTS; PR00907; THRBOMODULIN.
DR SMART; SM00032; CCP; 2.
DR SMART; SM00179; EGF_CA; 9.
DR SMART; SM00001; EGF_Like; 5.
DR PROSITE; PS00010; ASX_HYDROXYL; 8.
```

```
DR PROSITE; PS01186; EGF_2; 9.
DR PROSITE; PS01187; EGF_CA; 10.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
FT NON_TER 1394 1394
SQ SEQUENCE 1394 AA; 152269 MW; CD29380E3162F68A CRC64;
Query Match 13.2%; Score 386; DB 5; Length 1394;
Best Local Similarity 32.1%; Pred. No. 2.2e-22;
Matches 99; Conservative 33; Mismatches 110; Indels 66; Gaps 18;
QY 224 CTAPPGAVQG-----HWREAPGAW-----DCSVENGGEHACNALPG 261
DB 563 CLCPFGYALNDNHIVTSLNSSFITDSTSETPSAHTCLDIDECSLANGNSHFQNEPG 622
QY 262 APRCQCPAGALQADGRSCTASATQSC---NDLCEHFVCPNDPQSGVSCMCETGYRLAA 318
DB 623 GFQCACPLGYALSEDMRTC--QDIDECILDSNGQCSQLCL---NORGGFACACETGTFELTP 677
QY 319 DQHRCEVDVDCILPSPQPCRVNTQGGFECCHYPNYDLVGE--CVEPVDPC---FRAN 373
DB 678 DGFCAIDECSDYNGNSDIDICINLLGTHACACERGYELAKDKLSCLD-VDEAGLLSGG 736
QY 374 CEYQCQPLNQT-SYLCVCAEGFAPIPHEPHRCQFCNQTACPA--DCDPNTQASC----- 425
DB 737 CSHEC--INKAGTECGCPLGY--ILNDDGR-----SCSPALVGCPTGQTSADGCA 785
QY 426 -ECPEGYLLDGFCTIDIDEC--ENGPGSCGVCHNLPGTFECICGP-----DSALARHI 476
DB 786 IECNPGYTLGSDRDKVDIDECQKONGG--CSHRCSNTEGSKFCSCPPGYELDSQDKTCQDI 844
QY 477 GTDCDSGK 484
DB 845 -DECDQDK 851
RESULT 10
P87363 PRELIMINARY; PRT; 708 AA.
AC P87363;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Fibrillin-1 (Fragment).
GN FN1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN 1
RP SEQUENCE FROM N.A.
RX MEDLINE=20152896; PubMed=10691037;
RA Zhou G., Price C.E., Rosenquist T.H., Gadson P.F., Godfrey M.;
RT Partial cloning and sequencing of chick fibrillin-1 cDNA.;
RL In Vitro Cell. Dev. Biol. Anim. 36:19-25(2000).
DR EMBL; U88872; AAB48531.1; -.
DR HSSP; P07204; 2ADX.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF-Ca.
DR InterPro; IPR002212; Fibril-assoc.
DR InterPro; IPR001491; Thrbomodulin.
DR Pfam; PF00008; EGF; 13.
DR Pfam; PF00663; TB; 2.
DR PRINTS; PR00907; THRBOMODULIN.
DR SMART; SM00179; EGF_CA; 14.
DR PROSITE; PS00010; ASX_HYDROXYL; 13.
DR PROSITE; PS01186; EGF_2; 10.
DR PROSITE; PS01187; EGF_CA; 13.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
FT NON_TER 708 708
SQ SEQUENCE 708 AA; 76164 MW; C247271C1DF73361 CRC64;
```

Query Match 12.9%; Score 376; DB 13; Length 708;  
Best Local Similarity 31.8%; Pred. No. 5.6e-22;  
Matches 112; Conservative 26; Mismatches 108; Indels 106; Gaps 20;

Qy 244 DCSVNGGCEHACNAIPGAPROCPCAGALOADGRSCTASATQSCNDLCEHFCVNPND--- 300  
Db 223 EGSIMNGGCEHACNAIPGAPROCPCAGALOADGRSCTASATQSCNDLCEHFCVNPND--- 300  
Qy 301 -----DQGSYSYSCMCTGYRLAADQHRCEVDVDDCILEPSPC-PQRCVNTQGGFECHCYP 353  
Db 273 DGGQCTNIPGEYRCLCYDGFMAEDMKTCVDVNECDLHPNICLSGTCTCENTKGSFICHCDM 332  
Qy 354 NYDLVDGE--CVEPVDPC--FRANCEYQCOPLN-QTSYLCVCAEGFA-----PIP 398  
Db 333 GYSYSGKGTGCTD-INCEIGAHCNDRHACVCTNIPGSKSCSSGWIIGNIKCTDLDECS 391  
Qy 399 HEPHRCQFCNCTACPADCPNTOAS--CECPGYILDDGFICTDIDE-----CENG- 448  
Db 392 NGTHKCSPH-----ADC-KNTMGSYRLCKEGY-TGDGFTCTDLDECSNINLCENGO 442  
Qy 449 -----GF-----CS-----GVCHNLPCTFEICGPDPSALA 473  
Db 443 CLNAPGGYRCECDMGFLPSLDKACEDIDECSLPNICVYGTCHNLPGLFRCECEVGYELD 502  
Qy 474 RHIG-----TDCSGK-VDGGDSGSGEPSPPTPGSTLTPPAVLV 513  
Db 503 RSGNCTDNECADPTTCISGTCVNTAGSYTCECPDP-----FELNPTRVGCV 550

RESULT 11  
Q9WHU9  
ID Q9WHU9 PRELIMINARY; PRT; 2906 AA.  
AC Q9WHU9  
DT 01-NOV-1999 (Tremblrel. 12, Created)  
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)  
DE Fibrillin-2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99350231; PubMed=10419698;  
RA Yang Q., Ota K., Tian Y., Kumar A., Wada J., Kashiwara N., Wallner E.,  
RA Kanwar Y.S.;  
RT "Cloning of rat fibrillin-2 cDNA and its role in branching  
RT morphogenesis of embryonic lung."  
RL Dev. Biol. 212:229-242(1999).  
DR EMBL; AF135060; AAD34439.1; -  
DR HSSP; P35555; IEMN.  
DR InterPro; IPR002086; Aldehyde\_dehydr.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR001438; EGF\_II.  
DR InterPro; IPR002212; Fibril-assoc.  
DR Pfam; PF00008; EGF; 46.  
DR PRINTS; PF00683; TB; 9.  
DR PRINTS; PR00010; EGFBLD.  
DR SMART; SM00179; EGF\_CA; 42.  
DR SMART; SM00001; EGF\_like; 4.  
DR PROSITE; PS00070; ALDEHYDE\_DEHYDR\_CYS; UNKNOWN\_1.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 43.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_2.  
DR PROSITE; PS01186; EGF\_2; 36.  
DR PROSITE; PS01187; EGF\_CA; 43.  
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
SQ SEQUENCE 2906 AA; 313371 MW; 9EE64E727044EF58 CRC64;

Query Match 12.9%; Score 375; DB 11; Length 2906;  
Best Local Similarity 32.9%; Pred. No. 3.8e-21;  
Matches 102; Conservative 24; Mismatches 92; Indels 92; Gaps 18;

Qy 244 DCSVNGGCEHACNAIPGAPROCPCAGALOADGRSCTASATQSCNDLCEHFCVNPND--- 301  
Db 1237 ECMIMNGGCTQCTNSEGSEYSCSGYALMPDRSCA-----DIDE--CENNPDC 1286  
Qy 302 -----QPGSYSCMCTGYRLAADQHRCEVDVDDCILEPSPC-PQRCVNTQGGFECHCYP 353  
Db 1287 DGGQCTNIPGEYRCLCYDGFMAEDMKTCIDVNECDLHPNICMFGECENTKGSFICHQCL 1346  
Qy 354 NYDLVDGE--CVEPVDPC--FRANCEYQCOPLN-QTSYLCVCAEGFA-----PIP 398  
Db 1347 GYSYSGKGTGCTD-VDECEIGAHCNCDHMSCLNVPGSKSCREGWVGNGIKCIDLDECA 1405  
Qy 399 HEPHRCQFCNCTACPADCPNTOAS--CECPGYILDDGFICTDIDE-----CENG- 448  
Db 1406 NGTHQCSI-----NAQC-VNTPGYSYRCACSEGF-TGDGFTCTSDVDECAENINLCENGO 1456  
Qy 449 -----GF-----CS-----GVCHNLPCTFEICGPDPSALA 473  
Db 1457 CLNAPGGYRCECEMGFTPASDSRSCQDIDECFQNICVFTGTCNNLPGMFHCICDDGVLG 1516  
Qy 474 RHIG--TDCCD 481  
Db 1517 RTGGHCTDID 1526

RESULT 12  
O75095  
ID O75095 PRELIMINARY; PRT; 1246 AA.  
AC O75095  
DT 01-NOV-1998 (Tremblrel. 08, Created)  
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)  
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)  
DE MEGF6 (Fragment).  
GN MEGF6.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=BRIN;  
RX MEDLINE=98360089; PubMed=9693030;  
RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;  
RT "Identification of high-molecular-weight proteins with multiple EGF-  
RT like motifs by motif-trap screening."  
RL Genomics 51:27-34(1998).  
DR EMBL; AB011539; BAA32467.2; -  
FT NON\_TER 1  
SQ SEQUENCE 1246 AA; 130304 MW; BD8E70A0F6A2CB29 CRC64;

Query Match 12.7%; Score 369; DB 4; Length 1246;  
Best Local Similarity 32.6%; Pred. No. 4.5e-21;  
Matches 95; Conservative 27; Mismatches 97; Indels 72; Gaps 15;

Qy 238 EAPGAWDCSVNGGCEHACNAIPGAPROCPCAGALOADGRSCTASATQSC---NDLCEH 294  
Db 70 QRPDVDECRTHNGGCHRCVNTPGSYLCCKPGFRLHDSRTCL--AINSCALNGGCOH 127  
Qy 295 FCV-----PNDPQGS-----YSCMCTGYR 315  
Db 128 HCVQLTITRHCQCRPGFQLQEDGRHCVRRSPCANRNGSCMHRCQVVRGLARCECHVGQ 187  
Qy 316 LAADQHRCEVDVDDCILEPSPCQRCVNTQGGFECHCYPNIDL-VDG-EC-----VEPVDPC 369  
Db 188 LAADGKACEDVDECAAGLAQCAHGLCNTQGSFKCVCHAGYELGADGRQCYRIEMEIVNSC 247  
Qy 370 FRAN--CEYQCOPLNQTYS--LCVCAEGFAPIPHEPHRCQMF-----NOTACPADCD 418  
Db 248 EANNNGCASHGC---SHTSAGPLCTCPRGY-----ELDTQRTCIDVDCCADSPCCQOVCT 299  
Qy 419 PNTQA-SCECPGYILD-DGFICTDIDECENG-GFCSGVCHNLPCTFEIC 466

Db 300 NNPGGYECGYAGYRLSADGGCGEDYDECASSRGGEHHCNTLAGSFQCS 350

## RESULT 13

O88840  
ID O88840 PRELIMINARY; PRT; 3857 AA.  
AC O88840;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Mutant fibrillin-1.  
GN FBN1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP "Structure of the mutant fibrillin-1 gene in the tight skin (TSK) mouse.";  
RL DNA Res. 4:267-271(1997).  
DR EMBL; AF007248; AAC62317.1; -.  
DR HSP; P35555; IAPU.  
DR MGD; MGI:95489; Fbn1.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR002557; Chitin\_bind\_per.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR002212; Fibril-assoc.  
DR InterPro; IPR000822; znf\_C2H2.  
DR Pfam; PF00008; EGF; 64.  
DR Pfam; PF00683; TB; 12.  
DR SMART; SM00494; ChtBD2; 2.  
DR SMART; SM00179; EGF\_Ca; 60.  
DR SMART; SM00001; EGF\_like; 4.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 61.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_2.  
DR PROSITE; PS01186; EGF\_2; 50.  
DR PROSITE; PS01187; EGF\_Ca; 61.  
DR PROSITE; PS00028; ZINC\_FINGER\_C2H2\_1; UNKNOWN\_1.  
DR KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
SQ SEQUENCE 3857 AA; 418301 MW; 5BC0618BC527E04C CRC64;

Query Match 12.6%; Score 367; DB 11; Length 3857;  
Best Local Similarity 30.3%; Pred. No. 2.2e-20;  
Matches 118; Conservative 28; Mismatches 132; Indels 112; Gaps 21;  
QY 212 SAAYAPLGL-----QLMCTAPPAGVQGHWAREAPGAWDCSVENGGEHACNAIPGAPRC 265  
Db 2148 SANLCPHRCVNLIGKQACNFGYHPTDRLFCVDIDECINMGSCETFCNTSDGSYEC 2207  
QY 266 QCPAGALQADGRSCTASATQSCNDLCEHFCVNP-----DQPGYSVCMTGYRL 316  
Db 2208 SCQGFALPDRSCT-----DIDE--CEDPNICDGGQCTNPGEYRCICVDGFMA 2257  
QY 317 AADQHRCEVDVDCILEPSPC-PQRCVNTQGGFCHCYPNYDLVDGE--CPEPVDPC--FR 371  
Db 2258 SEDMKTCDVNECDLNPILSTCENTKSGFICHCDMGSGKKGTGCTD-INECEIGA 2316  
QY 372 ANCEYOCQPLNOT--SYLCVCAEGFA-----PIPEPHRCQMFNCQTACPADCPN 420  
Db 2317 HNCDRHAVCTNAGSFRCSCSPGWIGDKICTDLDECSSGTHMCSQH-----ADC-KN 2368  
QY 421 TQAS--CECEPEGLDDGFTCTDIDEC-EN-----GGF-----450  
Db 2369 TMGSYRCLCKDGY-TGDGFTCTDLDECSENLCNGQCINAFPGYRCBCDMGFVPSADG 2427  
QY 451 -----CS-----GVCHNLPGFECICGPDPSALAHIG-----TDCDSGK 484

Db 2428 KACEDIDECSLPNICVFCGTCHNLPGFPRCECEIGVELDRSGGNCCTDYNECLDPTTCISGN 2487

QY 485 -VDGDSGSGEPFPPSPTPGSTLTTPAVGLV 513

Db 2488 CVNTPGSGYTCDCPPD----FELNPTRVGCV 2513

## RESULT 14

Q96JJB8  
ID Q96JJB8 PRELIMINARY; PRT; 2809 AA.  
AC Q96JJB8;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Fibrillin3.  
GN KIAA1776.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RX MEDLINE=21245130; PubMed=11347906;  
RA Nagase T., Nakayama M., Nakajima D., Kikuno R., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XX.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 8:85-95(2001).  
DR EMBL; AB053450; BAB47408.1; -.  
DR InterPro; IPR000152; Asx\_hydroxyl.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR002212; Fibril-assoc.  
DR Pfam; PF00008; EGF; 45.  
DR Pfam; PF00683; TB; 9.  
DR PROSITE; PS00010; ASX\_HYDROXYL; UNKNOWN\_41.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_2.  
DR PROSITE; PS01186; EGF\_2; UNKNOWN\_36.  
DR PROSITE; PS01187; EGF\_Ca; UNKNOWN\_40.  
SQ SEQUENCE 2809 AA; 300323 MW; 20C04CC006C0161F CRC64;

Query Match 12.5%; Score 365; DB 4; Length 2809;  
Best Local Similarity 30.0%; Pred. No. 2.3e-20;  
Matches 107; Conservative 27; Mismatches 107; Indels 116; Gaps 20;  
QY 244 DCSVENGGCEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP--- 300  
Db 1158 ECRVQNGGCDVHCINTEGYSRCSCGGYSLMPDGRACA-----DYDE--CEENPRVC 1207  
QY 301 -----DQPGYSVCMTGYRLAADQHRCEVDVDCILEPSPCPO-RCVNTQGGFCHCY 353  
Db 1208 DQGHCTNMPGGHRCICVDGFMAIPDMRTCDVDDECDLPHICLHGDGENTKGSFVCHQL 1267  
QY 354 NYDLVDGE--CPEPVDPC-----FRANC-----374  
Db 1268 GYMWKRGATGCS-D-VDECEVGHNCDSHASCINIPGSPSCRLPGWGDGFECHDLDECI 1326  
QY 375 --EYQCP-----LN-QTSYLCVCAEGFAPIPHEPHRCQMF--NQTACPADCD--PNTQA- 423  
Db 1327 SQEHRCSPRGDCNLNPGYSRCTCROGA-----GDGFFCEDRDECAENVDCDNGOCL 1379  
QY 424 -----SECEPGY-ILDDGFTCTDIDECENGFGFS-GVCHNLPGFECICGPDPSALARH 475  
Db 1380 NAGPYRCBCEMGFDPTEDHRACQDWDCAQGNLCFAFSCENLPGMFRICNG-----1432  
QY 476 ICTDCDSKGVDSGSGEPFPPSP-----TPGS-----TLTPAVGLV 513  
Db 1433 -GYELDRG---GGNCTDINEADPNVNCINGVCINTPGSYLCSCPDGFELNPGVGCV 1485  
RESULT 15  
Q9WUHB

Search completed: December 16, 2002, 17:26:17  
Job time : 40.5 secs

```
ID Q9WUH8 PRELIMINARY; PRT; 2872 AA.
AC Q9WUH8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Fibrillin-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID-10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-95032689; PubMed-9815129;
RA Kanwar Y.S., Ota K., Yang Q., Kumar A., Wada J., Kashiwara N.,
RA Peterson D.R.;
RT *Isolation of rat fibrillin-1 cDNA and its relevance in metanephric
RT development.*;
RL Am. J. Physiol. 275:F710-F723(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Kanwar Y.S.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF135059; AAD34438.1; -.
DR HSP; P35555; IAP1.
DR InterPro; IPR000152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR001881; EGF_CA.
DR InterPro; IPR002212; Fibril-assoc.
DR InterPro; IPR000822; Znf_C2H2.
DR Pfam; PF00008; EGF; 45.
DR Pfam; PF00683; TB; 9.
DR SMART; SM00179; EGF_CA; 41.
DR SMART; SM00001; EGF_like; 5.
DR PROSITE; PS00010; ASX_HYDROXYL; 42.
DR PROSITE; PS00022; EGF_1; UNKNOWN_2.
DR PROSITE; PS01186; EGF_2; 38.
DR PROSITE; PS01187; EGF_CA; 41.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ SEQUENCE 2872 AA; 312069 MW; 0C4F7F3B87A80280 CRC64;

Query Match 12.4%; Score 361; DB 11; Length 2872;
Best Local Similarity 30.6%; Pred. NO. 4.9e-20;
Matches 114; Conservative 25; Mismatches 127; Indels 106; Caps 20;

QY 224 CTAPPGAVQGHWAEPAGWDCSVEGCGEHCAGNAIPGAPRCOCPAGAAQAQAGRSCTAS 283
DB 1181 CACNPGYHPTHDLFCVDIDECSIMNGCETCTNSDGYECSQCPGALMPQORSET-- 1238
QY 284 ATQSCNDLCHEFCVNPDP-----DQPGSYSCMCETGYRLAADOHRCEVDVDDCILEPS 334
DB 1239 -----DIDE--CEDNPNICDGGQCTNIPGEYRCLCYDGFWASEDMKTCVDVNECDLNP 1290
QY 335 PC-PQRCVNTQGGFECHCPNYDLVDGE--CPEVPDPC--FRANCEYQCQPLNOT-SYLC 388
DB 1291 ICLSGTCENTKGFICHCDMGYSCKKTKGCTD-INECEIGHNCGRHAVCTNTAGSFKC 1349
QY 389 VCAEGFA-----PIPEPHRCOMFONQACPADCPDPTOAS--CECPEGYILDDG 436
DB 1350 SCSPGWIGDGIKCTDLDSCNSGTHMCSQH-----ADC-KNTMGYSRCLCKDGY-TGDG 1400
QY 437 FICTDIDECE-----NG-----GF-----CS-----G 453
DB 1401 FTCTDLDSESNLNLNGQCLNAPAGYRCEDMGFVPSADGKACEDINECSLNPICVFG 1460
QY 454 VCHNLPGTFCICGPDSSALARHIG-----TDCDSGK-VDGGDSGSGEPPSPPTP 501
DB 1461 TCHNLPLGLFCECEIGYELDRSGGNCVDVNECLDPTTTCISGNCVNTPGSYTCDCPPD--- 1517
QY 502 GSTLTTPPAVGLV 513
DB 1518 -FELNPTRVGCV 1528
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:23:13 ; Search time 13.5 Seconds  
(without alignments)  
1124.609 Million cell updates/sec

Title: US-09-509-994-2

Perfect score: 2316  
Sequence: 1 MGVVLGALAGLGFAP.....PSPTPGSLTPRAVLVHSG 516

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	2916	100.0	575	1	US-08-312-870-1
2	2916	100.0	575	6	Sequence 1, Appli Patent No. 546668
3	2912	99.9	575	1	US-08-261-206A-59
4	2908	99.7	575	1	Sequence 59, Appl Patent No. 5256770
5	2846.5	97.6	572	6	US-08-170-290A-54
6	2826	96.9	498	2	Sequence 2, Appli Patent No. 5256770
7	2824	96.8	497	1	Sequence 3, Appli Patent No. 5256770
8	2766	94.9	494	1	US-08-312-870-3
9	2766	94.9	494	1	US-08-014-723-14
10	2764	94.8	494	1	US-08-110-011A-14
11	2764	94.8	494	1	Sequence 14, Appl US-08-014-723-16
12	2690	92.2	475	1	Sequence 16, Appl US-08-110-011A-16
13	2690	92.2	475	1	Sequence 2, Appli US-08-307-444A-2
14	2686	92.1	475	1	Sequence 2, Appli US-08-587-389-2
15	2686	92.1	475	1	Sequence 1, Appli US-08-307-444A-1
16	2686	92.1	475	1	Sequence 1, Appli US-08-587-389-1
17	2680	91.9	476	1	Sequence 1, Appli US-08-014-723-1
18	2678	91.8	476	1	Sequence 1, Appli US-08-110-011A-1
19	2678	91.8	476	1	Sequence 2, Appli US-08-014-723-2
20	2678	91.8	476	1	Sequence 18, Appl US-08-014-723-18
21	2678	91.8	476	1	Sequence 2, Appli US-08-110-011A-2
22	2596	89.0	456	1	Sequence 18, Appl US-08-110-011A-18
23	2596	89.0	456	1	Sequence 4, Appli US-08-307-444A-4
24	2592	88.9	456	1	Sequence 4, Appli US-08-587-389-4
25	2592	88.9	456	1	Sequence 3, Appli US-08-307-444A-3
26	2543	87.2	446	1	Sequence 3, Appli US-08-587-389-3
27	2543	87.2	446	1	Sequence 5, Appli US-08-307-444A-5
					Sequence 5, Appli US-08-587-389-5

28	1621	55.6	275	1	US-08-312-870-7	Sequence 7, Appli
29	1159	39.7	215	1	US-08-312-870-5	Sequence 5, Appli
30	689	23.6	115	1	US-08-312-870-9	Sequence 9, Appli
31	681	23.4	114	2	US-08-733-564-1	Sequence 1, Appli
32	584	20.0	652	2	US-08-751-305-2	Sequence 2, Appli
33	573	19.7	492	4	US-09-724-864-39	Sequence 39, Appl
34	356	12.2	638	2	US-08-897-443-1	Sequence 1, Appli
35	352	12.1	58	1	US-08-261-206A-3	Sequence 3, Appli
36	334.5	11.5	1964	4	US-09-467-997-1	Sequence 1, Appli
37	326.5	11.2	956	2	US-08-897-443-3	Sequence 3, Appli
38	323	11.1	1253	3	US-08-479-722B-4	Sequence 4, Appli
39	322.5	11.1	1394	6	5177197-30	Patent No. 5177197
40	319.5	11.0	1833	5	US-08-479-722B-2	Sequence 2, Appli
41	319.5	11.0	1833	5	PCT-US95-02251-18	Sequence 18, Appl
42	315.5	10.8	443	2	US-08-833-963C-2	Sequence 2, Appli
43	315.5	10.8	443	3	US-08-880-514-1	Sequence 1, Appli
44	308	10.6	448	2	US-08-884-072-1	Sequence 1, Appli
45	308	10.6	448	4	US-09-212-168-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-08-312-870-1  
; Sequence 1, Application US/08312870  
; Patent No. 5639625  
; GENERAL INFORMATION:  
; APPLICANT: Carson, Craig W.  
; APPLICANT: Esmon, Charles T.  
; TITLE OF INVENTION: Method for Detecting Antibodies to  
; TITLE OF INVENTION: Thrombomodulin in Patients  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Richards, Medlock & Andrews  
; STREET: 1201 Elm Street, Suite 4500  
; CITY: Dallas  
; STATE: Texas  
; COUNTRY: US  
; ZIP: 75270-2197  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/312,870  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hansen, Eugenia S.  
; REGISTRATION NUMBER: 31,966  
; REFERENCE/DOCKET NUMBER: OMRF B35150  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 214-939-4500  
; TELEFAX: 214-939-4600  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 575 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: N-terminal  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 19..575  
; US-08-312-870-1

Query Match 100.0%; Score 2916; DB 1; Length 575;  
Best Local Similarity 100.0%; Pred. No. 8.3e-199;  
Matches 516; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLGVLVLGALALAGLGFPAPEPQPGSQCVHDCFALYPGPATFLNASQICDGLRGHLM 60  
Db 1 MLGVLVLGALALAGLGFPAPEPQPGSQCVHDCFALYPGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVLSILLNGDGVGRRRLWIGLQPPGCGDKRGLGPGFQWVTGDNNTSYS 120  
Db 61 TVRSSVAADVLSILLNGDGVGRRRLWIGLQPPGCGDKRGLGPGFQWVTGDNNTSYS 120  
QY 121 RWARLDLNGAPLCPLCYAVSAAEATVPSEPIWEEQCEVKADGFCEHFFHFPATCRPLAV 180  
Db 121 RWARLDLNGAPLCPLCYAVSAAEATVPSEPIWEEQCEVKADGFCEHFFHFPATCRPLAV 180  
QY 181 EPGAAAAVSYITGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240  
Db 181 EPGAAAAVSYITGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240  
QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
Db 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DQPGSYSCMCEYGRYLAADQHRCEYDVCILPSPCQRCVNTQGGFECCHYFNVDLVDG 360  
Db 301 DQPGSYSCMCEYGRYLAADQHRCEYDVCILPSPCQRCVNTQGGFECCHYFNVDLVDG 360  
QY 361 ECVEPVDPCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPN 420  
Db 361 ECVEPVDPCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPN 420  
QY 421 TOASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALARHIGTDC 480  
Db 421 TOASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALARHIGTDC 480  
QY 481 DSGKVDGDSGSGEPPTPGSTLTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPTPGSTLTPPAVGLVHSG 516  
RESULT 2  
546668-6  
; Patent No. 546668  
; APPLICANT: GLASER, CHARLES B.; MORSE, MICHAEL J.; LIGHT,  
; DAVID R.  
; TITLE OF INVENTION: SUPERIOR THROMBOMODULIN ANALOGS FOR  
; PHARMACEUTICAL USE  
; NUMBER OF SEQUENCES: 57  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/155,346  
; FILING DATE: 22-NOV-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 568,456  
; FILING DATE: 15-AUG-1990  
; APPLICATION NUMBER: 506,325  
; FILING DATE: 09-APR-1990  
; APPLICATION NUMBER: 406,941  
; FILING DATE: 13-SEP-1989  
; APPLICATION NUMBER: 345,374  
; FILING DATE: 28-APR-1989  
; SEQ ID NO: 6  
; LENGTH: 575  
546668-6

Query Match 100.0%; Score 2916; DB 6; Length 575;  
Best Local Similarity 100.0%; Pred. No. 8.3e-199; Indels 0; Gaps 0;  
Matches 516; Conservative 0; Mismatches 0;

QY 1 MLGVLVLGALALAGLGFPAPEPQPGSQCVHDCFALYPGPATFLNASQICDGLRGHLM 60  
Db 1 MLGVLVLGALALAGLGFPAPEPQPGSQCVHDCFALYPGPATFLNASQICDGLRGHLM 60  
QY 61 TVRSSVAADVLSILLNGDGVGRRRLWIGLQPPGCGDKRGLGPGFQWVTGDNNTSYS 120  
Db 61 TVRSSVAADVLSILLNGDGVGRRRLWIGLQPPGCGDKRGLGPGFQWVTGDNNTSYS 120

QY 121 RWARLDLNGAPLCPLCYAVSAAEATVPSEPIWEEQCEVKADGFCEHFFHFPATCRPLAV 180  
Db 121 RWARLDLNGAPLCPLCYAVSAAEATVPSEPIWEEQCEVKADGFCEHFFHFPATCRPLAV 180  
QY 181 EPGAAAAVSYITGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240  
Db 181 EPGAAAAVSYITGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240  
QY 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
Db 241 GAWDCSVENGCGEHACNAIPGAPRCQCPAGALQADGRSCTASATQSCNDLCEHFCVNP 300  
QY 301 DQPGSYSCMCEYGRYLAADQHRCEYDVCILPSPCQRCVNTQGGFECCHYFNVDLVDG 360  
Db 301 DQPGSYSCMCEYGRYLAADQHRCEYDVCILPSPCQRCVNTQGGFECCHYFNVDLVDG 360  
QY 361 ECVEPVDPCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPN 420  
Db 361 ECVEPVDPCFRANCEYQCOPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDPN 420  
QY 421 TOASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALARHIGTDC 480  
Db 421 TOASCECEGYILDDGFICTDIDECENGFCGVCVCHNLPGTFECICGPDSSALARHIGTDC 480  
QY 481 DSGKVDGDSGSGEPPTPGSTLTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPTPGSTLTPPAVGLVHSG 516  
RESULT 3  
US-08-261-206A-59  
; Sequence 59, Application US/08261206A  
; Patent No. 5574007  
; GENERAL INFORMATION:  
; APPLICANT: Gomi, Mitichitaka  
; APPLICANT: Zushi, Komakazu  
; APPLICANT: Yamamoto, Shuji  
; APPLICANT: Suzuki, Koji  
; APPLICANT: Matsuda, Akio  
; TITLE OF INVENTION: A Polypeptide Capable of Interacting  
; TITLE OF INVENTION: with Thrombin  
; NUMBER OF SEQUENCES: 80  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Birch, Stewart, Kolasch & Birch  
; STREET: 301 N. Washington St.  
; CITY: Falls Church  
; STATE: Virginia  
; COUNTRY: USA  
; ZIP: 22046-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/261,206A  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/740,492  
; FILING DATE: 03-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Svensson, Leonard R.  
; REGISTRATION NUMBER: 30330  
; REFERENCE/DOCKET NUMBER: 216-275P  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 703-241-1300  
; TELEFAX: 703-241-2848  
; TELEX: 248345  
; INFORMATION FOR SEQ ID NO: 59:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 575 amino acids



TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..575  
OTHER INFORMATION: /label= protein  
OTHER INFORMATION: /note= "human thrombomodulin"  
US-08-261-206A-59

Query Match 99.9%; Score 2912; DB 1; Length 575;  
Best Local Similarity 99.8%; Pred. No. 1.6e-198;  
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGVLVLGALALAGLGFPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
Db 1 MGVLVLGALALAGLGFPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60

Qy 61 TVRSSVAADVLSLLNGDGGVRRRLWGLQLPPCGDPPKRLGPGFQWVGTGDNNTSYS 120  
Db 61 TVRSSVAADVLSLLNGDGGVRRRLWGLQLPPCGDPPKRLGPGFQWVGTGDNNTSYS 120

Qy 121 RWARDLNGAPLCGPLCVAVSAEATVPSEPIWEEQQCEVKADGFLCEFFHFPATCRPLAV 180  
Db 121 RWARDLNGAPLCGPLCVAVSAEATVPSEPIWEEQQCEVKADGFLCEFFHFPATCRPLAV 180

Qy 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVGHHAREAP 240  
Db 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVGHHAREAP 240

Qy 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFVCPNP 300  
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFVCPNP 300

Qy 301 DQGSYSVCMECYGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
Db 301 DQGSYSVCMECYGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360

Qy 361 ECVEPVDPFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
Db 361 ECVEPVDPFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420

Qy 421 TQASCECEGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDSSALARHIGTDC 480  
Db 421 TQASCECEGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDSSALARHIGTDC 480

Qy 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/170,290A  
FILING DATE: 28-DEC-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/05573  
FILING DATE: 01-JUL-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/724,237  
FILING DATE: 01-JUL-1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Heslin, James M.  
REGISTRATION NUMBER: 29,541  
REFERENCE/DOCKET NUMBER: 11972-58-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-326-2400  
TELEFAX: 415-326-2422  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 575 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-170-290A-54

Query Match 99.7%; Score 2908; DB 1; Length 575;  
Best Local Similarity 99.8%; Pred. No. 3.1e-198;  
Matches 515; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MGVLVLGALALAGLGFPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60  
Db 1 MGVLVLGALALAGLGFPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGHLM 60

Qy 61 TVRSSVAADVLSLLNGDGGVRRRLWGLQLPPCGDPPKRLGPGFQWVGTGDNNTSYS 120  
Db 61 TVRSSVAADVLSLLNGDGGVRRRLWGLQLPPCGDPPKRLGPGFQWVGTGDNNTSYS 120

Qy 121 RWARDLNGAPLCGPLCVAVSAEATVPSEPIWEEQQCEVKADGFLCEFFHFPATCRPLAV 180  
Db 121 RWARDLNGAPLCGPLCVAVSAEATVPSEPIWEEQQCEVKADGFLCEFFHFPATCRPLAV 180

Qy 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVGHHAREAP 240  
Db 181 EPGAAAAVSYTYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVAVGHHAREAP 240

Qy 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFVCPNP 300  
Db 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFVCPNP 300

Qy 301 DQGSYSVCMECYGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360  
Db 301 DQGSYSVCMECYGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360

Qy 361 ECVEPVDPFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
Db 361 ECVEPVDPFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420

Qy 421 TQASCECEGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDSSALARHIGTDC 480  
Db 421 TQASCECEGYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDSSALARHIGTDC 480

Qy 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516  
Db 481 DSGKVDGDSGSGEPPTPGSTLTTPPAVGLVHSG 516

RESULT 5

```
5256770-7
; Patent No. 5256770
; APPLICANT: GLASER, CHARLES B.; MORSER, MICHAEL J.; LIGHT,
; DAVID R.
; TITLE OF INVENTION: OXIDATION RESISTANT THROMBOMODULIN ANALOGS
; NUMBER OF SEQUENCES: 48
; CURRENT APPLICATION DATA:
; FILING DATE: 09-APR-1990
; APPLICATION NUMBER: US/07/506,325
; SEQ ID NO: 7:
; LENGTH: 572
5256770-7

Query Match          97.6%; Score 2846.5; DB 6; Length 572;
Best Local Similarity 98.6%; Pred. No. 6.8e-194;
Matches 509; Conservative 0; Mismatches 4; Indels 3; Gaps 2;

QY 1 MLGVILGALAGLGFPAEPPQPGSGQCVHDCFALYPGPAFLNASQICDGLRGLHM 60
DB 1 MLGVILGALAGLGFPAEPPQPGSGQCVHDCFALYPGPAFLNASQICDGLRGLHM 60
QY 61 TVRSSVAADVISLLNGDGVRRRLWIGLQPLPPGCGDPKRLGFGFQWVTGDNNTSYS 120
DB 61 TVRSSVAADVISLLNGDGVG-RRRLWIGLQPLPPGCGDPKRLGFGFQWVTGDNNTSYS 119
QY 121 RWARLDUNGAPLGLCPICVAVSAEATVPSEPIWEEQCEVKADGFLCEPHFPATCRPLAV 180
DB 120 RWARLDUNGAPLGLCPICVAVSAEATVPSEPIWEEQCEVKADGFLCEPHFPATCRPLAV 179
QY 181 EPGAAAVSITVTGTPFAARGADFOALPGVSSAAVAPLGLQMLCTAPPAGVGHWAREAP 240
DB 180 EPGAAAVSITVTGTPFAARGADFOALPGVSSAAVAPLGLQMLCTA--GNVOGHWAREAP 237
QY 241 GAWDCSVENGCGCEHACNAPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 300
DB 238 GAWDCSVENGCGCEHACNAPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNP 297
QY 301 DQPGSVSCMETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECVCPNYDLVDG 360
DB 298 DQPGSVSCMETGYRLAADQHRCEVDVDCILEPSPQRCVNTQGGFECVCPNYDLVDG 357
QY 361 ECVEPVDPCFRANCEYQCPPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDN 420
DB 358 ECVEPVDPCFRANCEYQCPPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDN 417
QY 421 TQASCEPEGYILDDGFICTDIDECENGCGFCVCHNLPGTFECICGPDSSALARHIGTDC 480
DB 418 TQASCEPEGYILDDGFICTDIDECENGCGFCVCHNLPGTFECICGPDSSALARHIGTDC 477
QY 481 DSGKVDGSDSGSEPPSPPTPGSTLTPPAVGLVHSG 516
DB 478 DSGKVDGSDSGSEPPSPPTPGSTLTPPAVGLVHSG 513

RESULT 6
US-08-733-564-2
; Sequence 2, Application US/08733564
; Patent No. 5916874
; GENERAL INFORMATION:
; APPLICANT: FUJIWARA, Kenji
; APPLICANT: MOCHIDA, Satoshi
; TITLE OF INVENTION: METHOD FOR TREATING LIVER INJURY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
; STREET: P.O. Box 747
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/733,564
FILING DATE: 18 OCTOBER 1996
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, Leonard R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 0216-0362P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 498 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Protein
US-08-733-564-2

Query Match          96.9%; Score 2826; DB 2; Length 498;
Best Local Similarity 99.8%; Pred. No. 1.7e-192;
Matches 497; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 19 APAEPQPGSGQCVHDCFALYPGPAFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 78
DB 1 APAEPQPGSGQCVHDCFALYPGPAFLNASQICDGLRGLHMTVRSSVAADVISLLNGD 60
QY 79 GGVRRRLWIGLQPLPPGCGDPKRLGFGFQWVTGDNNTSYSRWARLDUNGAPLGLCV 138
DB 61 GGVRRRLWIGLQPLPPGCGDPKRLGFGFQWVTGDNNTSYSRWARLDUNGAPLGLCV 120
QY 139 AVSAEATVPSEPIWEEQCEVKADGFLCEPHFPATCRPLAVEPGAAAVSITVTGPEA 198
DB 121 AVSAEATVPSEPIWEEQCEVKADGFLCEPHFPATCRPLAVEPGAAAVSITVTGPEA 180
QY 199 ARGADFOALPGVSSAAVAPLGLQMLCTAPPAGVGHWAREAPGAWDCSVENGCGCEHACNA 258
DB 181 ARGADFOALPGVSSAAVAPLGLQMLCTAPPAGVGHWAREAPGAWDCSVENGCGCEHACNA 240
QY 259 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYSCMETGYRLAA 318
DB 241 IPGAPRCQCPAGAAALQADGRSCTASATQSCNDLCEHFCVNPDPGYSYSCMETGYRLAA 300
QY 319 DQHRCEVDVDCILEPSPQRCVNTQGGFECVCPNYDLVDGCEVPEVDPFCFRANCEYQC 378
DB 301 DQHRCEVDVDCILEPSPQRCVNTQGGFECVCPNYDLVDGCEVPEVDPFCFRANCEYQC 360
QY 379 QPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDNPTQASCEPEGYILDDGFI 438
DB 361 QPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCOTACPADCDNPTQASCEPEGYILDDGFI 420
QY 439 CTDIDECENGCGFCVCHNLPGTFECICGPDSSALARHIGTDCDSDGKVDGSDSGSEPPSP 498
DB 421 CTDIDECENGCGFCVCHNLPGTFECICGPDSSALVRHIGTDCDSDGKVDGSDSGSEPPSP 480
QY 499 PTPGSTLTPPAVGLVHSG 516
DB 481 PTPGSTLTPPAVGLVHSG 498

RESULT 7
US-08-312-870-3
; Sequence 3, Application US/08312870
; Patent No. 5639625
; GENERAL INFORMATION:
; APPLICANT: Carson, Craig W.
; APPLICANT: Esmon, Charles T.
; TITLE OF INVENTION: Method for Detecting Antibodies to
; TITLE OF INVENTION: Thrombomodulin in Patients
; NUMBER OF SEQUENCES: 11
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richards, Medlock & Andrews
; STREET: 1201 Elm Street, Suite 4500
; CITY: Dallas
; STATE: Texas
; COUNTRY: US
; ZIP: 75270-2197
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/312,870
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hansen, Eugenia S.
; REGISTRATION NUMBER: 31,966
; REFERENCE/DOCKET NUMBER: OMRF B35150
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 214-939-4500
; TELEFAX: 214-939-4600
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 497 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-312-870-3

Query Match          96.8%; Score 2824; DB 1; Length 497;
Best Local Similarity 100.0%; Pred. No. 2.3e-192;
Matches 497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 APAEPQGGSCVDEHDFALYPGATFLNASQICDGLRGHLMTVRSSVAADVISLLNGD 78
DB 1 APAEPQGGSCVDEHDFALYPGATFLNASQICDGLRGHLMTVRSSVAADVISLLNGD 60
QY 79 GVGRRRLWIGLQPLPGCGDKRGLRGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 138
DB 61 GVGRRRLWIGLQPLPGCGDKRGLRGFQWVTGDNNTSYSRWARLDLNGAPLCGLCV 120
QY 139 AVSAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAAASITYGTFFA 198
DB 121 AVSAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAVEPGAAAAAASITYGTFFA 180
QY 199 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCCBHACNA 258
DB 181 ARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAPGAWDCSVENGCCBHACNA 240
QY 259 IPGAPRCQCPAGALQADGRSCTASATOSCNLDCEHFCVPPNPDPQPGSYSCMCETGYRLAA 318
DB 241 IPGAPRCQCPAGALQADGRSCTASATOSCNLDCEHFCVPPNPDPQPGSYSCMCETGYRLAA 300
QY 319 DOHRCEDVDDCILEPSPQRCVNTQGGFECYCHYNYDLVDGECVPEVDPDCFRANCEYQC 378
DB 301 DOHRCEDVDDCILEPSPQRCVNTQGGFECYCHYNYDLVDGECVPEVDPDCFRANCEYQC 360
QY 379 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQNTACPDQNTQASCECEGYILLDDGFI 438
DB 361 QPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQNTACPDQNTQASCECEGYILLDDGFI 420
QY 439 CTDIDECENGFGCSGVCHNLPGTFTFCICGPDSSALARIHGTDCDSKVGDDSGSGSEPPPS 498
DB 421 CTDIDECENGFGCSGVCHNLPGTFTFCICGPDSSALARIHGTDCDSKVGDDSGSGSEPPPS 480
QY 499 PTPGSTLTTPPAGVLVHS 515
DB 481 PTPGSTLTTPPAGVLVHS 497
```

RESULT 8

```
US-08-014-723-14
; Sequence 14, Application US/08014723
; Patent No. 5273962
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,723
; FILING DATE: 19930208
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5273962man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-071-0 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-014-723-14

Query Match          94.9%; Score 2766; DB 1; Length 494;
Best Local Similarity 99.2%; Pred. No. 2.9e-188;
Matches 490; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLGVLVLGALALAGLGFPAPEPQPGSQCVDEHDFALYPGATFLNASQICDGLRGHLM 60
DB 1 MLGVLVLGALALAGLGFPAPEPQPGSQCVDEHDFALYPGATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVISLLNGDGGVGRRRRLWIGLQPLPGCGDKRGLRGFQWVTGDNNTSY 120
DB 61 TVRSSVAADVISLLNGDGGVGRRRRLWIGLQPLPGCGDKRGLRGFQWVTGDNNTSY 120
QY 121 RWARLDLNGAPLCGLCVAVSAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180
DB 121 RWARLDLNGAPLCGLCVAVSAEATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180
QY 181 EPGAAAAAASITYGTFFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
DB 181 EPGAAAAAASITYGTFFAARGADFOALPVGSSAAVAPLGLQMLCTAPPGAVQGHWAREAP 240
QY 241 GAWDCSVENGCCBHACNAIPGAPRCQCPAGALQADGRSCTASATOSCNLDCEHFCVPPNP 300
DB 241 GAWDCSVENGCCBHACNAIPGAPRCQCPAGALQADGRSCTASATOSCNLDCEHFCVPPNP 300
QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDDCILEPSPQRCVNTQGGFECYCHYNYDLVDG 360
DB 301 DQPGSYSCMCETGYRLAADQHRCEVDVDDCILEPSPQRCVNTQGGFECYCHYNYDLVDG 360
QY 361 ECVEPVDPDCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQNTACPDQDN 420
```

```
|||||
Db 361 ECVEPDPFCFRANCEYQCPLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN 420
QY 421 TQASCECEGYLDGDFICTDIDECENGFGCGVCHNLPGTFECICGPDSSALARHIGTDC 480
Db 421 TQASCECEGYLDGDFICTDIDECENGFGCGVCHNLPGTFECICGPDSSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
Db 481 DSGKVDGDSGSGE 494

RESULT 9
US-08-110-011A-14
; Sequence 14, Application US/08110011A
; Patent No. 5354664
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,011A
; FILING DATE: 23-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5354664man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-073-0 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-110-011A-14

Query Match 94.9%; Score 2766; DB 1; Length 494;
Best Local Similarity 99.2%; Pred. No. 2.9e-188;
Matches 490; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLGVLVLGALALAGLGFAPAEPPQGGSCQVEHDCFPALYPGPATFLNASQICDGLRGLHLM 60
Db 1 MLGVLVLGALALAGLGFAPAEPPQGGSCQVEHDCFPALYPGPATFLNASQICDGLRGLHLM 60
QY 61 TVRSSVAADVVISLLNGDCGVRRLRWIGLQLPPCGDPRKRLGPLRGFWVTGDNNTSYS 120
Db 61 TVRSSVAADVVISLLNGDCGVRRLRWIGLQLPPCGDPRKRLGPLRGFWVTGDNNTSYS 120
QY 121 RWRDLNLGAPLCGPLCAVSAEAATVPSEPTWEEQQCEVKADGFTCEHFFPATCPEPLAV 180
Db 121 RWRDLNLGAPLCGPLCAVSAEAATVPSEPTWEEQQCEVKADGFTCEHFFPATCPEPLAV 180
```

```
QY 181 EPGAAAAVSTYCTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
Db 181 EPGAAAAVSTYCTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
QY 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
Db 241 GAWDCSVENGCGEHCACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300
QY 301 DPGSYSCMCETGYRLAADOHRCEDVDCCILEPSPCPORCVNTQGGFCHCYPNYDLVDG 360
Db 301 DPGSYSCMCETGYRLAADOHRCEDVDCCILEPSPCPORCVNTQGGFCHCYPNYDLVDG 360
QY 361 ECVEPDPFCFRANCEYQCPLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN 420
Db 361 ECVEPDPFCFRANCEYQCPLNQTSYLCVCAEGFAPIPHEPHRCOMFCNQACPADCDPN 420
QY 421 TQASCECEGYLDGDFICTDIDECENGFGCGVCHNLPGTFECICGPDSSALARHIGTDC 480
Db 421 TQASCECEGYLDGDFICTDIDECENGFGCGVCHNLPGTFECICGPDSSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
Db 481 DSGKVDGDSGSGE 494

RESULT 10
US-08-014-723-16
; Sequence 16, Application US/08014723
; Patent No. 5273962
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,723
; FILING DATE: 19930208
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5273962man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-071-0 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000
; TELEFAX: (703)413-2220
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-014-723-16

Query Match 94.8%; Score 2764; DB 1; Length 494;
Best Local Similarity 99.0%; Pred. No. 4e-188;
Matches 489; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

```

QY 1 MLGVLVIGALALAGLGFPPAPAEPOPGSGOCVEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MLGVLVIGALALAGLGFPPAPAEPOPGSGOCVEHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPKRLGRLGFGQWVTGDNNTSYS 120
DB 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPKRLGRLGFGQWVTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGLPLCAVSAAEATVPSEPIWEEQOCEVKADGFLCEHFFHFPATCRPLAV 180
DB 121 RWARLDLNGAPLCGLPLCAVSAAEATVPSEPIWEEQOCEVKADGFLCEHFFHFPATCRPLAV 180
QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
QY 241 GAWDCSVENGCGEHACNATPGAPRCOCAGAAQADGRSCTASATOSCNLCEHFCVNP 300
DB 241 GAWDCSVENGCGEHACNATPGAPRCOCAGAAQADGRSCTASATOSCNLCEHFCVNP 300
QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
DB 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
QY 361 ECVEPVDPCFRANCEYQCOPLNQTSLVCVAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
DB 361 ECVEPVDPCFRANCEYQCOPLNQTSLVCVAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
QY 421 TQASCEPGEYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDSSALVRHIGTDC 480
DB 421 TQASCEPGEYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDSSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
DB 481 DSGKVDDEASGSGD 494

RESULT 11
US-08-110-011A-16
; Sequence 16, Application US/08110011A
; Patent No. 5354664
; GENERAL INFORMATION:
; APPLICANT: Doi, Takeshi
; APPLICANT: Iwasaki, Akio
; APPLICANT: Saino, Yushi
; APPLICANT: Kimura, Shigeru
; APPLICANT: Ohkuchi, Masao
; TITLE OF INVENTION: Thrombin-Binding Substance and Process
; TITLE OF INVENTION: For Preparing the Same
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/110,011A
; FILING DATE: 23-AUG-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5354664man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 80-073-0 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)413-3000

```

```

; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-110-011A-16

Query Match 94.8%; Score 2764; DB 1; Length 494;
Best Local Similarity 99.0%; Pred. No. 4e-188;
Matches 489; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MLGVLVIGALALAGLGFPPAPAEPOPGSGOCVEHDCFALYPGPATFLNASQICDGLRGHLM 60
DB 1 MLGVLVIGALALAGLGFPPAPAEPOPGSGOCVEHDCFALYPGPATFLNASQICDGLRGHLM 60
QY 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPKRLGRLGFGQWVTGDNNTSYS 120
DB 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQPLPGCGDPKRLGRLGFGQWVTGDNNTSYS 120
QY 121 RWARLDLNGAPLCGLPLCAVSAAEATVPSEPIWEEQOCEVKADGFLCEHFFHFPATCRPLAV 180
DB 121 RWARLDLNGAPLCGLPLCAVSAAEATVPSEPIWEEQOCEVKADGFLCEHFFHFPATCRPLAV 180
QY 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
DB 181 EPGAAAASVITYGTPFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP 240
QY 241 GAWDCSVENGCGEHACNATPGAPRCOCAGAAQADGRSCTASATOSCNLCEHFCVNP 300
DB 241 GAWDCSVENGCGEHACNATPGAPRCOCAGAAQADGRSCTASATOSCNLCEHFCVNP 300
QY 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
DB 301 DQPGSYSCMCTGYRLAADOHRCEDVDDCILEPSPQRCVNTQGGFECCHYPNYDLVDG 360
QY 361 ECVEPVDPCFRANCEYQCOPLNQTSLVCVAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
DB 361 ECVEPVDPCFRANCEYQCOPLNQTSLVCVAEGFAPIPHEPHRCOMFCNOTACPADCDPN 420
QY 421 TQASCEPGEYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDSSALVRHIGTDC 480
DB 421 TQASCEPGEYILDDGFICTDIDECENGFCGVCNHLPGTFECICGPDSSALVRHIGTDC 480
QY 481 DSGKVDGDSGSGE 494
DB 481 DSGKVDDEASGSGD 494

RESULT 12
US-08-307-444A-2
; Sequence 2, Application US/08307444A
; Patent No. 5516659
; GENERAL INFORMATION:
; APPLICANT: NII, ATSUSHI
; APPLICANT: MORISHITA, HIDEAKI
; APPLICANT: UEMURA, AKIO
; APPLICANT: MOCHIDA, EI
; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OLIFF & BERRIDGE
; STREET: P.O. BOX 19928
; CITY: ALEXANDRIA
; STATE: VA
; COUNTRY: USA
; ZIP: 22320
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```

;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/307,444A  
;; FILING DATE: 19-SEP-1994  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 07/835,436  
;; FILING DATE: 26-FEB-1992  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OLIFF, JAMES A.  
;; REFERENCE/DOCKET NUMBER: JAO 27706  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703) 836-6400  
;; TELEFAX: (703) 836-2787  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 475 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;;  
US-08-307-444A-2

Query Match 92.2%; Score 2690; DB 1; Length 475;  
Best Local Similarity 100.0%; Pred. No. 6.5e-183;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MLGVLVLGALAGLGPAPAEPPQGGSCQVEHDCFPATFLNASQICDGLRGLHM 60  
DB 1 MLGVLVLGALAGLGPAPAEPPQGGSCQVEHDCFPATFLNASQICDGLRGLHM 60  
  
QY 61 TVRSSVAADVLSILLNGDGGVRRRLWIGLQPLPGCGDKRGLRGLRQFQWVTGDNNTSYS 120  
DB 61 TVRSSVAADVLSILLNGDGGVRRRLWIGLQPLPGCGDKRGLRGLRQFQWVTGDNNTSYS 120  
  
QY 121 RWARLDNLGAPLCGLPCVAVSAEAATVPSEPIWEEQCEVKADGFLCEHFFPACRPLAV 180  
DB 121 RWARLDNLGAPLCGLPCVAVSAEAATVPSEPIWEEQCEVKADGFLCEHFFPACRPLAV 180  
  
QY 181 EPGAAAASVITYGTPTFAARGADFPALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240  
DB 181 EPGAAAASVITYGTPTFAARGADFPALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240  
  
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEHFCVNP 300  
DB 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEHFCVNP 300  
  
QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECCHYCNPDLDVVG 360  
DB 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECCHYCNPDLDVVG 360  
  
QY 361 ECVEPVDPFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
DB 361 ECVEPVDPFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
  
QY 421 TQASCEPEGYILDDGFTCTDIDBENGFGFCGVCNHLNPTFTFECICGPDSSALARRH 475  
DB 421 TQASCEPEGYILDDGFTCTDIDBENGFGFCGVCNHLNPTFTFECICGPDSSALARRH 475

RESULT 13  
US-08-587-389-2  
; Sequence 2, Application US/08587389  
; Patent No. 5695964  
; GENERAL INFORMATION:  
; APPLICANT: NII, ATSUSHI  
; APPLICANT: MORISHITA, HIDEAKI  
; APPLICANT: UEMURA, AKIO  
; APPLICANT: MOCHIDA, EI

;; TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT  
;; TITLE OF INVENTION: PRODUCTION THEREOF, AND THERAPEUTIC AGENT (AS AMENDED)  
;; NUMBER OF SEQUENCES: 27  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: OLIFF & BERRIDGE  
;; STREET: P.O. BOX 19928  
;; CITY: ALEXANDRIA  
;; STATE: VA  
;; COUNTRY: USA  
;; ZIP: 22320  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC Compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/587,389  
;; FILING DATE: 17-JAN-1996  
;; CLASSIFICATION: 435  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/307,444  
;; FILING DATE: 19-SEP-1994  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: OLIFF, JAMES A.  
;; REFERENCE/DOCKET NUMBER: JAO 27706  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (703) 836-6400  
;; TELEFAX: (703) 836-2787  
;; TELEX: 90-1799 PTO ALEX  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 475 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; HYPOTHETICAL: NO  
;; ANTI-SENSE: NO  
;;  
US-08-587-389-2

Query Match 92.2%; Score 2690; DB 1; Length 475;  
Best Local Similarity 100.0%; Pred. No. 6.5e-183;  
Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MLGVLVLGALAGLGPAPAEPPQGGSCQVEHDCFPATFLNASQICDGLRGLHM 60  
DB 1 MLGVLVLGALAGLGPAPAEPPQGGSCQVEHDCFPATFLNASQICDGLRGLHM 60  
  
QY 61 TVRSSVAADVLSILLNGDGGVRRRLWIGLQPLPGCGDKRGLRGLRQFQWVTGDNNTSYS 120  
DB 61 TVRSSVAADVLSILLNGDGGVRRRLWIGLQPLPGCGDKRGLRGLRQFQWVTGDNNTSYS 120  
  
QY 121 RWARLDNLGAPLCGLPCVAVSAEAATVPSEPIWEEQCEVKADGFLCEHFFPACRPLAV 180  
DB 121 RWARLDNLGAPLCGLPCVAVSAEAATVPSEPIWEEQCEVKADGFLCEHFFPACRPLAV 180  
  
QY 181 EPGAAAASVITYGTPTFAARGADFPALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240  
DB 181 EPGAAAASVITYGTPTFAARGADFPALPVGSSAAVAPLGLQMLCTAPPAGVQGHWAREAP 240  
  
QY 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEHFCVNP 300  
DB 241 GAWDCSVENGCGCEHACNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDLCHEHFCVNP 300  
  
QY 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECCHYCNPDLDVVG 360  
DB 301 DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPCQRCVNTQGGFECCHYCNPDLDVVG 360  
  
QY 361 ECVEPVDPFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420  
DB 361 ECVEPVDPFCFRANCEYQCQPLNQTSLYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCDPN 420

QY 421 TOASCCEPGYILDDGFICTDIDCEGNGFCGVCVCHNLPCTFEICIGPDSALARH 475  
 Db 421 TOASCCEPGYILDDGFICTDIDCEGNGFCGVCVCHNLPCTFEICIGPDSALARH 475

RESULT 14  
 US-08-307-444A-1  
 ; Sequence 1, Application US/08307444A  
 ; Patent No. 5516659  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NII, ATSUSHI  
 ; APPLICANT: MORISHITA, HIDEAKI  
 ; APPLICANT: UEMURA, AKIO  
 ; APPLICANT: MOCHIDA, EI  
 ; TITLE OF INVENTION: ANTICOAGULANT POLYPEPTIDES  
 ; NUMBER OF SEQUENCES: 27  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OLIFF & BERRIDGE  
 ; STREET: P.O. BOX 19928  
 ; CITY: ALEXANDRIA  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22320

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/307,444A  
 FILING DATE: 19-SEP-1994  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/835,436  
 FILING DATE: 26-FEB-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: OLIFF, JAMES A.  
 REGISTRATION NUMBER: 27,075  
 REFERENCE/DOCKET NUMBER: JAO 27706  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 836-6400  
 TELEFAX: (703) 836-2787  
 TELEX: 90-1799 PTO ALEX  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 475 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO

US-08-307-444A-1  
 Query Match 92.1%; Score 2686; DB 1; Length 475;  
 Best Local Similarity 99.8%; Pred. No. 1.3e-182;  
 Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MLGVLVLGALAGLGPAPAEPPQPGSQCVHEHDCFCALYPGPATFLNASQICDGLRHLH 60  
 Db 1 MLGVLVLGALAGLGPAPAEPPQPGSQCVHEHDCFCALYPGPATFLNASQICDGLRHLH 60  
 QY 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQLPPCGGDKPKRLGFGQWVTGDNNTSYS 120  
 Db 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQLPPCGGDKPKRLGFGQWVTGDNNTSYS 120  
 QY 121 RWARLDLNGAPLCPLCAVSAATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
 Db 121 RWARLDLNGAPLCPLCAVSAATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
 QY 181 EPGAAAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVGGVGHWAREAP 240  
 Db 181 EPGAAAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVGGVGHWAREAP 240

US-08-307-444A-1  
 Query Match 92.1%; Score 2686; DB 1; Length 475;  
 Best Local Similarity 99.8%; Pred. No. 1.3e-182;  
 Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MLGVLVLGALAGLGPAPAEPPQPGSQCVHEHDCFCALYPGPATFLNASQICDGLRHLH 60  
 Db 1 MLGVLVLGALAGLGPAPAEPPQPGSQCVHEHDCFCALYPGPATFLNASQICDGLRHLH 60  
 QY 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQLPPCGGDKPKRLGFGQWVTGDNNTSYS 120  
 Db 61 TVRSSVAADVLSLLNGDGGVRRRLWIGLQLPPCGGDKPKRLGFGQWVTGDNNTSYS 120  
 QY 121 RWARLDLNGAPLCPLCAVSAATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
 Db 121 RWARLDLNGAPLCPLCAVSAATVPSEPIWEEQCEVKADGFLCEHFPATCRPLAV 180  
 QY 181 EPGAAAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVGGVGHWAREAP 240  
 Db 181 EPGAAAAVSTYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPVGGVGHWAREAP 240

QY 241 GAWDCSVENGCGEHCACNAIPGAPRCOCPPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
 Db 241 GAWDCSVENGCGEHCACNAIPGAPRCOCPPAGAAQADGRSCTASATQSCNDLCEHFCVNP 300  
 QY 301 DQPGSYSCMCFGYRLAADOHRCEDVDDCILBSPCPQRCVANTOGGFECHCYPNYDLVDG 360  
 Db 301 DQPGSYSCMCFGYRLAADOHRCEDVDDCILBSPCPQRCVANTOGGFECHCYPNYDLVDG 360  
 QY 361 ECVEPVDPCCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCQMFQACPDADCPN 420  
 Db 361 ECVEPVDPCCFRANCEYQCOPLNOTSYLCVCAEGFAPIPHEPHRCQMFQACPDADCPN 420  
 QY 421 TOASCCEPGYILDDGFICTDIDCEGNGFCGVCVCHNLPCTFEICIGPDSALARH 475  
 Db 421 TOASCCEPGYILDDGFICTDIDCEGNGFCGVCVCHNLPCTFEICIGPDSALVRH 475

RESULT 15  
 US-08-587-389-1  
 ; Sequence 1, Application US/08587389  
 ; Patent No. 5655964  
 ; GENERAL INFORMATION:  
 ; APPLICANT: NII, ATSUSHI  
 ; APPLICANT: MORISHITA, HIDEAKI  
 ; APPLICANT: UEMURA, AKIO  
 ; APPLICANT: MOCHIDA, EI  
 ; TITLE OF INVENTION: TRUNCATED THROMBOMODULIN, RECOMBINANT  
 ; NUMBER OF SEQUENCES: 27  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OLIFF & BERRIDGE  
 ; STREET: P.O. BOX 19928  
 ; CITY: ALEXANDRIA  
 ; STATE: VA  
 ; COUNTRY: USA  
 ; ZIP: 22320

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/587,389  
 FILING DATE: 17-JAN-1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/307,444  
 FILING DATE: 19-SEP-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: OLIFF, JAMES A.  
 REGISTRATION NUMBER: 27,075  
 REFERENCE/DOCKET NUMBER: JAO 27706  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 836-6400  
 TELEFAX: (703) 836-2787  
 TELEX: 90-1799 PTO ALEX  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 475 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 US-08-587-389-1

Query Match 92.1%; Score 2686; DB 1; Length 475;  
 Best Local Similarity 99.8%; Pred. No. 1.3e-182;  
 Matches 474; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MLGVLVLGALAGLGPAPAEPPQPGSQCVHEHDCFCALYPGPATFLNASQICDGLRHLH 60

Db	1	MLGVVLGALALAGLGFPAPEPQPGSGOCVEHDCFALYPGPATFLNASQICDGLRGHLM	60
Qy	61	TVSSVAAADVISLLNGDGVGRRRLWIGLQPPGCGDKRLGRLRGFOWVTGDNNTSYS	120
Db	61	TVSSVAAADVISLLNGDGVGRRRLWIGLQPPGCGDKRLGRLRGFOWVTGDNNTSYS	120
Qy	121	RWARLDLNGAPLCPLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEPHFPATCRPLAV	180
Db	121	RWARLDLNGAPLCPLCVAVSAAEATVPSEPIWEEQOCEVKADGFLCEPHFPATCRPLAV	180
Qy	181	EPGAAAAVSIYGTTPFAARGADFQALPYGSSAAVAPLGLQLMCTAPPGAVOGHWAREAP	240
Db	181	EPGAAAAVSIYGTTPFAARGADFQALPYGSSAAVAPLGLQLMCTAPPGAVOGHWAREAP	240
Qy	241	GAWDCSVENGGBHACNAIPGAPRCOCPAGALQADGRSCTASATQSCNDLCEHFCVNP	300
Db	241	GAWDCSVENGGBHACNAIPGAPRCOCPAGALQADGRSCTASATQSCNDLCEHFCVNP	300
Qy	301	DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQSGFECHCYPNYDLVDG	360
Db	301	DQPGSYSCMCETGYRLAADQHRCEVDVDCILEPSPQRCVNTQSGFECHCYPNYDLVDG	360
Qy	361	ECVEPVDPCFRANCEYQCQPLNQTSLVCVCAEGFADIPHEPHRCOMFCNQACPADCDPN	420
Db	361	ECVEPVDPCFRANCEYQCQPLNQTSLVCVCAEGFADIPHEPHRCOMFCNQACPADCDPN	420
Qy	421	TQASCECEGYIILDDGFICTDIDECENGFGCSGVCHNLPGTPECICGPDALARH	475
Db	421	TQASCECEGYIILDDGFICTDIDECENGFGCSGVCHNLPGTPECICGPDALVRH	475

Search completed: December 16, 2002, 17:27:18  
Job time : 15.5 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 16, 2002, 17:26:23 ; Search time 8.5 Seconds  
(without alignments)  
1011.502 Million cell updates/sec

Title: US-09-509-994-2  
Perfect score: 2916  
Sequence: 1 MLGVLVLAGALAGLGFAP.....PSPTGSLTPPAVLVHSG 516

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105981 seqs, 1662342 residues  
Total number of hits satisfying chosen parameters: 105981

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : Published\_Applications\_AA.\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2912	99.9	575	10	US-09-938-405-2
2	585	20.1	652	10	US-09-789-919-96
3	547	18.8	644	10	US-09-789-919-62
4	547	18.8	644	10	US-09-789-919-73
5	353	12.1	776	9	US-10-000-512-8
6	353	12.1	959	9	US-10-000-512-10
7	352.5	12.1	877	10	US-09-764-898-200
8	348.5	12.0	915	9	US-09-905-291A-34
9	348.5	12.0	915	10	US-09-909-320-34
10	348.5	12.0	915	10	US-09-909-088B-34
11	332	11.4	999	10	US-09-747-371-2
12	327.5	11.2	974	10	US-09-898-570-14
13	327.5	11.2	1009	10	US-09-898-570-16
14	325	11.1	800	10	US-09-764-853-800
15	325	11.1	800	10	US-09-764-898-272
16	323	11.1	534	10	US-09-804-156-14
17	323	11.1	534	10	US-09-946-633-6
18	314	10.8	997	10	US-09-747-371-3
19	313.5	10.8	845	10	US-09-898-570-12

ALIGNMENTS

RESULT 1  
US-09-938-405-2  
; Sequence 2, Application US/09938405  
; Patent No. US20020111296A1  
; GENERAL INFORMATION:  
; APPLICANT: Festoff, Barry W.  
; APPLICANT: Morser, Michael J.  
; TITLE OF INVENTION: Thrombomodulin Analogs for Use in Recovery of Spinal Cord Injury  
; FILE REFERENCE: 51960AUSMI  
; CURRENT APPLICATION NUMBER: US/09/938,405  
; CURRENT FILING DATE: 2001-08-23  
; PRIOR APPLICATION NUMBER: 60/229,714  
; PRIOR FILING DATE: 2000-08-31  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: Patent version 3.1  
; SEQ ID NO 2  
; LENGTH: 575  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-938-405-2

Query Match	99.9%;	Score	2912;	DB	10;	Length	575;
Best Local Similarity	99.8%;	Pred. No.	1.3e-174;				
Matches	515;	Conservative	0;	Mismatches	1;	Indels	0;
Gaps	0;						
QY	1	MLGVLVLGALALAGLGFAPAPAPQPGSGCQVHDCFPALPGPATFLNASQICDGLRGLHM	60				
DB	1	MLGVLVLGALALAGLGFAPAPAPQPGSGCQVHDCFPALPGPATFLNASQICDGLRGLHM	60				
QY	61	TYRSSVAADVISILLNGDGGVGRRLWIGLQPPGCGDKPRLGPLRGFQWVTGDNNTSYS	120				
DB	61	TYRSSVAADVISILLNGDGGVGRRLWIGLQPPGCGDKPRLGPLRGFQWVTGDNNTSYS	120				
QY	121	RWARDLNCAPLCGPLCVASAAEATVPSEPIWEEQOCVKADGFCFHFATCPLAV	180				
DB	121	RWARDLNCAPLCGPLCVASAAEATVPSEPIWEEQOCVKADGFCFHFATCPLAV	180				
QY	181	EPGAAAAAVSITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP	240				
DB	181	EPGAAAAAVSITYGTPTFAARGADFOALPVGSSAAVAPLGLQLMCTAPPAGVQGHWAREAP	240				
QY	241	GAWDCSVENGGEHACNALIPGAPRCQC	PAGAAALQADGRSCTASATSCNDLCEHFCVPNP	300			

```
Db 241 GAWDCSVENGCGEACNAIPCAPRCQCPAGAAQADGSCSTASATQSCNDLCEHFCVNP 300
QY 301 DQPSYSCMCTGYRLAADHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360
Db 301 DQPSYSCMCTGYRLAADHRCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDG 360
QY 361 ECVPEVDPFRANCEYOCPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCPN 420
Db 361 ECVPEVDPFRANCEYOCPLNOTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCPN 420
QY 421 TQASCEPEGYIILDDGICTDIDCENGGFCGVCCHNLPGTFECICGPDSSALARHIGTDC 480
Db 421 TQASCEPEGYIILDDGICTDIDCENGGFCGVCCHNLPGTFECICGPDSSALARHIGTDC 480
QY 481 DSGKVDGDSGSGRPPSPPTPGSTLTTPPAVGLVHSG 516
Db 481 DSGKVDGDSGSGRPPSPPTPGSTLTTPPAVGLVHSG 516

RESULT 2
US-09-789-919-96
; Sequence 96, Application US/09789919
; Patent No. US20020064855A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor
; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM
; FILE REFERENCE: 2275-1-005
; CURRENT APPLICATION NUMBER: US/09/789,919
; CURRENT FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 96
; LENGTH: 652
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-789-919-96

Query Match 20.1%; Score 585; DB 10; Length 652;
Best Local Similarity 31.8%; Pred. No. 1.3e-29;
Matches 181; Conservative 56; Mismatches 209; Indels 124; Gaps 32;

QY 2 LGVLVLCALALA--GLGEPAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLH 59
Db 5 MGLLLLLLLLLLTPQAGTGADTE----AVVCGTACTYAHSGKLSAAEAQHNCNNGGNL 60
QY 60 MTRVSSVAAD----VISLLINGDGGVGR--RLWIGLQLPPG--CGDPKRLGLRGLRQFWVT 112
Db 61 ATVKSEEAQHVRVLAQLLRREAALTARMSKFWIGLQREKCKLDPSL--PLKGSFWG 118
QY 113 GDNNTSYSRWARLDLNGALPCGLPLCA--VSAEAATVPSE--PIWEEOQCEV-----EVKAD 163
Db 119 GGEDTPYSNNHKELRNSC--ISKRCVSLLDLSQLPLNRLPKWSGPGCGSPGSGSNTIE 176
QY 164 GFLCEHFPATCRPLAV--EPGAAAASVITYTTPFAARGADFOALPVGSSAAVAPLGLQL 222
Db 177 GFVKFSFSGMCRALGALGPG-----QVYTTTPQTSSLSLEAVPFASAANVA----- 224
QY 223 MC-TAPPAGVQGHW---AREAPGAWD-----CSVENGCGCHAC-NAIPGAP 263
Db 225 -CGEGDKDETQSHYFLCKEKAQPDVDMGSSGPLCVSPKYGCNFGNNGGCHQDCFEQDGSF 283
QY 264 RQCCPAGAAQADGSCSTASATQSCNDLCE--HFCVNPDPQCGSYSCMCTGYRLAADQH 321
Db 284 LCGRPGRELLDDLYTC-ASRNPSSPCRGATCVLGP-HGKNTICRCPQGYQLDSSQL 341
QY 322 RCEDVDDCILEPSPQRCVNTQGGFECHCYPNYDLVDGECVEPDPFRANCEYOCQPL 381
Db 342 DCVDVDEC--QDSPCAQECVNTPGFRCEWVG-----EPGPG-----GEGACQDV 386
QY 382 NCTSYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCPNTQAS--CECPEGYIL--DDGF 437
```

```
Db 387 DE-----CALGRSP-----CAQGC--TNTDGSFHCSEEGYVLAGEDGT 423
QY 438 ICTDIDEC--ENGFCGSGVCHNLPGTFECICGPDSSALARHIGTDCDSKV-----D 486
Db 424 QCQDVDECVGPGGLDCLCFNTQGSFHCGLPLGVLAPN-GVSCMTGMPVSLGPPSPGPD 482
QY 487 GGDGSGSG-----PPSPPTPGSTLTTPPA 509
Db 483 EEDKGEREGSVTPRAATASPTGPEGTPKA 512

RESULT 3
US-09-789-919-62
; Sequence 62, Application US/09789919
; Patent No. US20020064855A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor
; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM
; FILE REFERENCE: 2275-1-005
; CURRENT APPLICATION NUMBER: US/09/789,919
; CURRENT FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 62
; LENGTH: 644
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-789-919-62

Query Match 18.8%; Score 547; DB 10; Length 644;
Best Local Similarity 30.2%; Pred. No. 2.9e-27;
Matches 168; Conservative 55; Mismatches 210; Indels 124; Gaps 27;

QY 5 LVLGALALAGLGFAPAEPPQGGSCQVEHDCFALYPGPATFLNASQICDGLRGLHMTVRS 64
Db 9 LLIGLGLQGPWAG--AAADSO--AVVCEGTACTYAHWGKLSAAEAQHRCNENGGNLATVKS 64
QY 65 SVAA-----DVISLLINGDGGVGR--RLWIGLQLPPGCGPKRLGLRGLRQFWTGDNNTS 118
Db 65 EEARHVQQAQALTQLLTKAPLEAKMGKFWIGLQREKNCNTYHDL--PMRGFSWVGGEDTA 123
QY 119 YSRWARLDLNGALPCGLPLCAVSAEAATVPSE--PIWEEOQCEV-----KADGFLCEHF 171
Db 124 YSNWYKASKSSCIFKRCVSLILDLSLTPHPSHLPKWHESPCTGPEAPGNSIEGFLCKNF 183
QY 172 PATCRPLAV--EPGAAAASVITYTTPFAARGADFOALPVGSSAAVAPLGLQ-----LM 223
Db 184 KGMCRPLALGGPG-----RVYTTTPQTSSLSLEAVPPASVANVA--CGDEAKSETHYFL 236
QY 224 CT-APPGAQVQHWARAP---GAWDCSVENGCGCHAC-NAIPGAPRCQCPAGAAQADG 277
Db 237 CNEKTGFI--HWGSSGPLCVSPKFGCSFNNGGCGDCEFGDGSFRCGCRPFRLLDDL 294
QY 278 RSCASATQSCNDLCEHFCVNPDPQPG-----SYSCMCTGYRLAADQHRCEDVD 327
Db 295 VTCAS-----RNPSSNPCTGGMCHSVPLSENITCRCPGSQLDSSQVHCVDID 344
QY 328 DCILEPSPQRCVNTQGGFECHCYPNYDLVDG---ECVEPVPDPFRANCEYOCQPLNQT 384
Db 345 EC--QDSPCAQCVNTLGSFHCCEWVGQ--PSGPKREACEDVDECAANSP----- 392
QY 385 SYLCVCAEGFAPIPHEPHRCQMFNCQTACPADCPNTQAS--CECPEGYIL--DDGFICT 440
Db 393 -----CAQGC-----NTDGSFYCSCKEGYIVSGEDSTQCE 423
QY 441 DIDECEEN--GGFCGSGVCHNLPGTFECICGPDSSALARHIGTDCDSG-----KVD 486
Db 424 DIDECSDARGNCDLSLCFNTDGSFRCGCPPGWELAPN-GVFCSRGTVPFSELPARPQKED 482
QY 487 GGDGSGSGEPPSPPTGKS 503
```

Db 483 NDRKRSTMPPTMPSS 499

RESULT 4

US-09-789-919-73

; Sequence 73, Application US/09789919

; Patent No. US20020064855A1

; GENERAL INFORMATION:

; APPLICANT: Lemischka, Ihor

; APPLICANT: Moore, Kateri

; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM

; FILE OF INVENTION: CELLS AND USES THEREOF

; FILE REFERENCE: 2275-1-005

; CURRENT APPLICATION NUMBER: US/09/789,919

; CURRENT FILING DATE: 2001-02-21

; NUMBER OF SEQ ID NOS: 96

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 73

; LENGTH: 644

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-789-919-73

Query Match 18.8%; Score 547; DB 10; Length 644;

Best Local Similarity 30.2%; Pred. No. 2.9e-27;

Matches 168; Conservative 55; Mismatches 210; Indels 124; Gaps 27;

QY 5 LVGLALAGLGPAPAEPOGGSCQVEHDCFALYGPATFLNASICDGLRGLMTVRS 64

Db 9 LLGLLGLQPWAG--AAADSQ--AVVCEGTACTYTAHMKLSAAEAQHRNENGGLATVKS 64

QY 65 SVAA---DVISILLNGDGVGR--RLWIGLQLPCCGDPKRLGRLGFWQWTGDNNTS 118

Db 65 EEEHVVQALTKKAPLEAKMKFWIGLQREKNGCTYHDL--PMRGFSWVGCGEDTA 123

QY 119 YSWARLDLNGALPGCLPVASAAEATPSE--PIWEEQCEV-----KADGFLCBFHF 171

Db 124 YSNWYKAKSSCFKRCVSLILDLSTPHSLPKWHESPCGTPPEAGNSTEGFLCKFNF 183

QY 172 PATCRPLAV--EPGAAAAVSIITYGTPFAARGADFOALPVGSSAAVAPLGLQ-----LM 223

Db 184 KMCRLALGGPG-----RVITYTPQATTSLEAVPFASVANVA--CGDEAKSETHYFL 236

QY 224 CT-APPGAVGVHAREAP-----GAWDCSVENGCGEHAC--NAIPGAPRCQCPAGAAALQADG 277

Db 237 CNEKTPGIF--HWGSSGPLCVSPKFGCSFNNGCQDCFEFGDGSFRGCRPGFRLLDDL 294

QY 278 RSTASATQSCNDLCEHFCVNPDPQG-----SYSCMCETGYRLAADOHRCEVD 327

Db 295 VTCAS-----RNPCCSNPCTGGGMCHSVPLSENITCRCPGSGYQLDSSQVHCVDID 344

QY 328 DCILERSPCQRCVNTGGFECCHCYPNYDLVDG---ECVEPVPDPCFRANCEYOCQPLNQT 384

Db 345 EC--QDSFCAQDCVNTLGSFCECWGYQ--PSGPKKEACEDVDCAANSR-----392

QY 385 SYLCVCAEGFAPIPHEPHRCQMFNCNATACDPCDNTQAS--CECPEGYIL--DDGFTCT 440

Db 393 -----CAQGC-----NTDGSFYCSCKEYIVSGEDSTQCE 423

QY 441 DIDECEN--GGFCSGVCHNLPGTEFCICGPDALAHRTGTDGSC-----KVD 486

Db 424 DIDECSDARGNCPDLSLNTDGSFRGCPGWELAPN--GVFCRSRTVFSFELPARPPQKED 482

QY 487 GDSGSGEPPTSPG 503

Db 483 NDRKRSTMPPTMPSS 499

RESULT 5

US-10-000-512-8

; Sequence 8, Application US/10000512

; Patent No. US20020164699A1

; GENERAL INFORMATION:

; APPLICANT: Shimkets, Richard A

; APPLICANT: Fernandes, Elma

; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES ENCODING SAME

; FILE REFERENCE: 15966-556

; CURRENT APPLICATION NUMBER: US/10/000,512

; CURRENT FILING DATE: 2001-10-23

; PRIOR APPLICATION NUMBER: 09/619,252

; PRIOR FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 60/167,785

; PRIOR FILING DATE: 1999-11-29

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 8

; LENGTH: 776

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-000-512-8

Query Match 12.1%; Score 353; DB 9; Length 776;

Best Local Similarity 31.2%; Pred. No. 4.3e-15;

Matches 85; Conservative 41; Mismatches 102; Indels 44; Gaps 15;

QY 245 CSVENGGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTA--SATQSCNDLCEHFCVNPDPQ 303

Db 283 CAMEDHNCCEQLCVNVPGSFVCECYSGYALAEGRKRVADVYCASENHGCEHCV--NAD-- 339

QY 304 GSYSCMCETGYRLAADOHRCEVDVDCILEPSPCQRCVNTGGFECCHCYPNYDL--VDGEC 362

Db 340 GSYLCQCHGEFALNDEKTCIKIDYCASSNHGCQYECVNTDSDSYSCHLKGFTLNPKKT 399

QY 363 VEPVDPGF--RANCEYOCQPLNQTSYLCVCAEGFAPIPH-----EPHRCQMF 408

Db 400 CRRINYCALNKPGEHECVNMEE--SYRCRHRGYTLDPNGRKPCSRVDHCAQDGHCEQLC 458

QY 409 NQTACPADCDNTQAS--CECPEGYILDDGF--ICTDIDEC---ENGFCSGVCHNLPGTF 462

Db 459 -----LNTEDSFVQCQSEGLINEDLKTSRVDYCLLSDHG--CEYSCVNDRSF 506

QY 463 ECICGPDALAHRTGTDGSCGVKVDG---GDSG 491

Db 507 ACQC--PEGHVLRSDGKTC--AKLDSALGADHG 535

RESULT 6

US-10-000-512-10

; Sequence 10, Application US/10000512

; Patent No. US20020164699A1

; GENERAL INFORMATION:

; APPLICANT: Shimkets, Richard A

; APPLICANT: Fernandes, Elma

; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES ENCODING SAME

; FILE REFERENCE: 15966-556

; CURRENT APPLICATION NUMBER: US/10/000,512

; CURRENT FILING DATE: 2001-10-23

; PRIOR APPLICATION NUMBER: 09/619,252

; PRIOR FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 60/167,785

; PRIOR FILING DATE: 1999-11-29

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 10

; LENGTH: 959

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-000-512-10

Query Match 12.1%; Score 353; DB 9; Length 959;

Best Local Similarity 31.2%; Pred. No. 5.2e-15;

Matches 85; Conservative 41; Mismatches 102; Indels 44; Gaps 15;

QY 245 CSVENGGCGEHACNAIPGAPRCQCPAGAAALQADGRSCTA--SATQSCNDLCEHFCVNPDPQ 303





APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/909,088B  
CURRENT FILING DATE: 2001-07-18  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 34  
LENGTH: 915  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-909-088B-34

Query Match 12.0%; Score 348.5; DB 10; Length 915;  
Best Local Similarity 34.3%; Pred. No. 9.5e-15;  
Matches 85; Conservative 36; Mismatches 108; Indels 19; Gaps 14;

QY 245 GSVENGCCERACNAIPGAPRCQCPAGAAQADGRSCTA-SATQSCNDLCEHFVCPVNPDPQ 303  
DB 283 CAMEDHNCQELCVNVPQSGVFCQCYSGYALAECDKRCVAVDYCASENHGCHECV-NAD-- 339

QY 304 GSYSCMCTGYRLAADHRCEDVDCCILEPSPQRCVNTQGGFECHCYPNYDL-VDGEC 362  
DB 340 GSYLCQCHGEGFALNDEKTCRINYCALNKPGECHCVNNEESYICRCHGYTLDPNGKT 399

QY 363 VEPVDPGCFRAN--CEYQCPN-OTSILVCYCAEGFAPIPHEPRCOM--PC--NOTACPA 415  
DB 400 CSRDVHCAQQDHGCEJLC--LNTEDSFVQCSEGL-INEDLTCSRDVYCLLSHDGCEY 456

QY 416 DC-DPNTQASCECEGYIL-DDGFCITDIDECENGGF-CSGVCHNLPPTGTFECICGPDAL 472  
DB 457 SCVNDRSFACQCEGHVLRSDGKTCALDSCALGDHGCHECHSVSSSDSFVQCQF-EGYI 515

QY 473 ARHIGTDC 480  
DB 516 LREDGKTC 523

RESULT 11  
US-09-747-371-2  
Sequence 2, Application US/09747371  
Patent No. US20020006616A1  
GENERAL INFORMATION:  
APPLICANT: Gish, Kurt  
APPLICANT: Mack, David  
TITLE OF INVENTION: No. US20020006616A1 Methods of Diagnosing Breast Cancer, Com  
FILE OF INVENTION: Screening for Breast Cancer Modulators  
FILE REFERENCE: A-69028/DJB/JJD  
CURRENT APPLICATION NUMBER: US/09/747,371  
CURRENT FILING DATE: 2000-12-21  
PRIOR APPLICATION NUMBER: PCT/ US/00/06952  
PRIOR FILING DATE: 2000-03-15  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 2  
LENGTH: 999  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-747-371-2

Query Match 11.4%; Score 332; DB 10; Length 999;  
Best Local Similarity 26.8%; Pred. No. 1.1e-13;  
Matches 101; Conservative 31; Mismatches 117; Indels 128; Gaps 18;

QY 182 PGAAAAVSITYGTPFAARGADFOALPVGSSAAVAPLGLQMLCTAPPAGVQGHWARE--- 238  
DB 9 PGAAWAVLILLLLP-----PL-LLLAGAVPPGGRGAAGQEDVD 47

QY 239 --APGAWDCSVENGCGEHA---CNAIPGAPRCQCPAGAAQADGRSCTASATQSCNDL-- 291  
DB 48 ECAQGLDGC-----HADALCQNTPTSYKCSCKPG--YQEGEQC-EDIDECGNELNG 96

QY 292 -CEHFCVPNPDPGYSVCMCTGYRLAADHRCEDVDCCILEPSPQRCVNTQGGFECH 350  
DB 97 GCVHDCI---NIPGNVRCCTGDFGFMHLADHGNCLDVEDLENNGCGQHTCVNVMSYBCC 153

QY 351 CYPNYDLVDGE--CVEPVD---PCFRAN--CEYQCPNQTSLYLCVCAEGF----- 394  
DB 154 CKEGFFLSDNQHTCIHRSEGLSKMKNKHGCHSHICKAPRGVACVACRPGFELAKNQD 213

QY 395 -APIPHEPHRCOMFCNOTA----- 412  
DB 214 ILTCNHGNGCGQSCDDTADGPECSCHPQYKMHDTGRSCLEREDVLEVTESNTTSVVDG 273

QY 413 -----CPADC-DPNTQASCECEGYILD-DGFCITDIDEC--NGG 449  
DB 274 DKRVKRRLLMETCAVNNGGCDRTCKDTSIGVHCSPGVGFTLQDQKTKDIDECOTRNG 333

QY 450 FCSGVCHNLPPTGTFECIC 466  
DB 334 -CDHFCKNIVGSDGCG 349

RESULT 12  
US-09-898-570-14  
Sequence 14, Application US/09898570  
Patent No. US20020123612A1  
GENERAL INFORMATION:  
APPLICANT: GERLACH, VALERIE L.  
APPLICANT: ELLERMAN, KAREN  
APPLICANT: MACDOUGALL, JOHN R.  
APPLICANT: SMITHSON, GLENDA  
TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND  
FILE OF INVENTION: METHODS OF USING THE SAME  
FILE REFERENCE: 15966-776CIP  
CURRENT APPLICATION NUMBER: US/09/898,570  
CURRENT FILING DATE: 2001-07-03  
PRIOR APPLICATION NUMBER: 60/198,293  
PRIOR FILING DATE: 2000-04-19  
PRIOR APPLICATION NUMBER: 60/198,645

```

RESULT 14
US-09-764-853-800
; Sequence 800, Application US/09764853
; Patent No. US20020090672A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

```

```

RESULT 13
US-09-898-570-16
; Sequence 16, Application US/09898570
; Patent No. US20020123612A1
; GENERAL INFORMATION:
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDDA
; TITLE OF INVENTION: NOVEL HUMAN PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND
; TITLE OF INVENTION: METHODS OF USING THE SAME
; FILE REFERENCE: 15966-776CIP

```

```

; FILE REFERENCE: PJ206
; CURRENT APPLICATION NUMBER: US/09/764,853
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 939
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 800
; LENGTH: 800
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (197)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (201)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (687)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-853-800

Query Match          11.1%; Score 325; DB 10; Length 800;
Best Local Similarity 31.6%; Pred. No. 2.4e-13;
Matches 86; Conservative 36; Mismatches 88; Indels 62; Gaps 18

QY 244 DCSVNGGGEHACNAIPGRCQCPAGAAQAQDRSC-----TASATQSCNDLC 292
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 478 EAAGN-PCSHSCHNMGTYCYSCPKGITIAAGERTCQIDECALGRHTCHAGQDCDNTI 536
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 293 EHFCEVPNPDPQSGSYSCM-CETGYRLAADQHRCEVDVDDCILEPSPQRQVNTQGGFECH 350
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 537 -----GSYRCVVRGSGFRTSDGLSCQDINEC-QESSPCQRCFNAIGSFHCG 584
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 351 CYPNYDLVGECVEYDPDFRANC--EYQOPLNQTSYLCV--CAEGFAPIPH----- 399
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 585 CEPGYOLKGRKCMD-VNECRQNVCPDQHCKN-TRGGYKCIDLCPNGMTKAENGTCIDID 642
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 400 -----EPRHRCMFCNCTACPADCPNTOAS--CEGPEGVILDD-GFICTDIDECEN-GGFC 451
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 643 ECKDGTHQERY--NQI-----CE-NTRGSYRCVCPRGYSQGVGRPCMDINEXEQVKPC 694
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 452 SGVCHNLPGTFFECICGPDASALARHI---GTDC 480
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 695 AHQCSNTPGCFKICIPP-----GOHLLGDKGKSC 722
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 15
US-09-764-898-272
; Sequence 272, Application US/09764898
; Patent No. US20020090673A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PJ201
; CURRENT APPLICATION NUMBER: US/09/764,898
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 311
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 272
; LENGTH: 800
; TYPE: prt
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (197)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (201)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (687)

```